dedicated to the messengers of change

and the power of their wisdom

MOLECULAR GENETICS OF THE LIPOPROTEIN GENES:

ASSOCIATIONS WITH HYPERLIPOPROTEINEMIA

AND NIDDM IN MICRONESIANS

by Wendy Ann Baker

A thesis submitted for the degree of Doctor of Philosophy

of The Australian National University

September 1991

STATEMENT

This thesis describes the results of a research project carried out under the supervision of Prof. S.W. Serjeantson, within the Human Genetics Group at the Australian National University, from May 1988 to September 1991.

The experiments and data analysis presented in this thesis are my own original work, except where otherwise acknowledged. The work described has not previously been submitted for a degree at this or any other University.

Wendy Ann Baker

ACKNOWLEDGEMENTS

This thesis has been made possible by the professional assistance and personal support of many people. Those acknowledged here are owed a special debt of gratitude, which I hope is repaid in part by a recognition of their contribution to the completion of this thesis.

I would like to acknowledge the donation of probes for this work from Dr D. Drayna, Genentech, Ca, USA, Dr M.C. Schotz, Lipid Research, Wadsworth Medical Center, UCLA, CA, USA, and Dr S.K. Karathanasis, Dept. Pediatrics, Harvard Medical School, Boston, MA, USA.

I would particularly like to thank Prof. Sue Serjeantson for her ability to merge an understanding of my nature with the exacting task of thesis supervision. Her guidance has been valuable and will be remembered. I also thank her for helping me to recognise the importance of forests as well as trees.

Dr Sue Wilson has provided crucial and appropriate statistical advice and Dr Simon Easteal has critically proofed the three results sections of the thesis. Thanks are due to both for their time and encouragement.

I am grateful to Dr Paul Zimmet, Dr Gary Dowse, Caroline Finch and Veronica Collins of the World Health Organization Collaborating Centre for the Epidemiology of Diabetes Mellitus, Royal Southern Memorial Hospital, Melbourne, for their collaboration and cooperation in providing me with

access to the blood samples and the databases of both the 1982 and 1987 Nauruan Surveys.

I would like to thank Dr Elizabeth Jazwinska and Brett White for passing on to me their knowledge of laboratory techniques and making me welcome during the initial stages of my Ph.D. Without the management skills of Marj Coggan and Lynn Croft the execution of experimental work presented in this thesis would have been much more time consuming.

Marj Coggan has tirelessly and meticulously proofed the majority of this thesis and is thanked along with Dr Phil Board, Juleen Cavanaugh, and Tom Howe who proofed the remaining sections.

A special message of thanks to the Witches of East Wing, Marj, Juleen and Veronica, and to the Cyrenes of Asherah for their companionship, harmony, sanity, and wit. I take this opportunity to express my appreciation to the people who have supported me on a personal level over the past three years. Susan Buller, Juleen Cavanaugh, Aileen Howe, Jane Symonds and Sue Weston have, in their own very individual ways, provided me with a network of care and spectrum of experiences which I carry with me permanently.

ABSTRACT

The apolipoprotein genes are important protein components of plasma lipoprotein fractions, act as cofactors for many of the lipoprotein enzymes, and interact with lipoprotein receptors. Defects in the genes for apolipoproteins, lipoprotein enzymes and lipoprotein receptors are now recognised as the cause of a number of clinical lipoproteinemias, as well as being identified as determinants of plasma cholesterol and triglyceride concentrations in the general population. The genes coding for apolipoproteins, lipoprotein enzymes and receptors have also been implicated in the aetiology of cardiovascular disease.

The primary metabolic alteration and underlying biochemical cause of non-insulin dependent diabetes mellitus remains undetermined. The disorder is often accompanied by general obesity and central adiposity, and these three states are associated with altered lipoprotein metabolism. The Micronesian population of the Republic of Nauru is affected by high prevalences of both non-insulin dependent diabetes and obesity and has formed the study population for this thesis. This study examines the contribution of genetic variability at eight lipoprotein loci to plasma cholesterol and triglyceride variability, central adiposity, general obesity, and to the occurrence of non-insulin dependent diabetes mellitus.

The frequencies of seven restriction fragment length polymorphisms, and of the common apolipoprotein E (APOE) alleles were established in the Micronesian population and these frequencies compared with those observed in other ethnic groups. *Msp*I polymorphisms at the apolipoprotein A1 (APOA1) and apolipoprotein A2 (APOA2) loci, *Taq*I polymorphisms at the apolipoprotein C2 (APOC2) and apolipoprotein D (APOD) loci, a *Dra*I polymorphism at the apolipoprotein C1 (APOC1) locus and *Pvu*II polymorphisms at the lipoprotein lipase (LPL) and low density lipoprotein receptor (LDLR) loci were all present in the population. Frequencies for the presence of the restriction sites ranged from 0.230 ± 0.035 to 0.839 ± 0.027. Micronesian APOE allele frequencies were APOE*2 0.048 ± 0.016, APOE*3 0.827 ± 0.029 and APOE*4 0.125 ± 0.026.

In general, the lipoprotein allele frequencies were closer to those reported for Oriental populations, and most dissimilar to those found in Caucasoid and African populations. Probe/enzyme combinations generated systems with polymorphic information contents ranging from 0.234 to 0.374. The restriction fragment length polymorphisms were "highly informative", given the restriction on information content due to diallelism, and the apolipoprotein E polymorphism was found to be "reasonably informative" (polymorphic information content 0.273).

The frequencies of these eight polymorphisms were examined in relation to diabetes, body mass index, waist-to-hip

ratio, plasma triglyceride level and plasma cholesterol level. The nature of associations discovered between these lipoprotein genes and the variables listed were examined in both univariate and multivariate analyses. The presence of the APOC2 *Taq*I 3.8kb allele was significantly associated with lowered plasma triglyceride levels and waist-to-hip ratios. The APOD *Taq*I 2.7kb allele was significantly associated with non-insulin dependent diabetes mellitus and was more common in individuals with high cholesterol levels and high waist-to-hip ratios. The primary associations in the population are believed to be between APOC2 and plasma triglyceride levels, and between APOD and non-insulin dependent diabetes. Limited evidence was provided for an association between the APOE*2 allele and non-insulin dependent diabetes mellitus, however, there was no support for the often reported associations between APOE*2 and hypertriglyceridemia, or APOE*4 and hypercholesterolemia.

Linkage analyses were performed between the eight lipoprotein loci under investigation and diabetic liability, hypercholesterolemia and hypertriglyceridemia. These analyses were performed under a range of assumptions for mode of inheritance. Close Linkage to these three disorders was excluded, under a model for dominant inheritance, for several of the loci examined. APOC2 and APOD, in particular, were excluded from linkage with hypertriglyceridemia, hypercholesterolemia and non-insulin dependent diabetes mellitus, under dominant inheritance. Moreover, APOC2 was excluded from linkage with

hypertriglyceridemia, and APOD from linkage with hypercholesterolemia, when these disorders were modelled under recessive inheritance.

The lack of power of single markers to detect linkage with complex disorders has been discussed and it is concluded that the use of either highly variable polymorphisms or genetic maps may be required if detection of the underlying genetic cause, or causes, of these disorders are to be established. Genetic analyses of these complex disorders are hindered by variable age of onset and confounding secondary metabolic effects which make the nomination of individuals to affected or unaffected status difficult. Furthermore an improved understanding of the interactions between obesity and central adiposity, and of their effects on glucose and lipoprotein metabolism and insulin action will improve the likelihood of identifying the genes responsible for establishing a background genetic predisposition to diabetes, obesity and hyperlipo-proteinemia.

PUBLICATIONS

Baker, W.A., G. Dowse, P. Zimmet and S.W. Serjeantson. 1991. Population and pedigree analysis in Micronesians of eight loci contributing to lipid metabolism in NIDDM. In *Proceedings of the 14th International Diabetes Federation Congress*. Excerpta Medica, Amsterdam. in press.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

SECTION 1

GENERAL INTRODUCTION

1.1 INTRODUCTION

Altered lipoprotein metabolism is strongly implicated in the aetiology of atherosclerosis (Goldstein and Brown, 1983; Hegele and Breslow, 1987; Breslow, 1988). In rare cases, simple monogenic disorders of apolipoproteins (Zannis and Breslow, 1984; Galton, 1985), the low density lipoprotein receptor (LDLR) (Goldstein *et al.*, 1982; Russell *et al.*, 1986; Falus and Romics, 1988) and lipoprotein lipase (Devlin *et al.*, 1990; Monsalve *et al.*, 1990) result in hyperlipoproteinemias. More commonly, single genetic determinants cannot be identified for the lipoproteinemias. In these situations it is likely that the disorders are polygenic or multifactorial in origin with the most probable candidate genes being those associated with lipid transport and metabolism, that is, the apolipoprotein genes, the genes coding for lipoprotein enzymes and the lipoprotein receptor genes (Galton, 1987).

Lipoprotein metabolism is altered in non-insulin dependent diabetes (NIDDM) and this alteration is accompanied by an increased risk for cardiac heart disease (CHD) (Brunzell *et al.*, 1985; Howard, 1987; Assmann and Schulte, 1988). The possibility exists that the apolipoprotein and lipoprotein receptor and enzyme genes are also involved in the aetiology of NIDDM.

1.2 MOLECULAR GENETICS OF THE LIPOPROTEIN GENES

Lipoproteins are the major carriers of triglycerides, cholesterol, cholesteryl esters and phospholipids in human

plasma. They are synthesized mainly in the liver and intestine and are complex macromolecular aggregates with cores of non-polar lipids and coats of relatively polar material (Smith *et al.*, 1978). The plasma lipoproteins vary greatly in size, lipid composition and peptide content. They have traditionally been separated by ultracentrifugal flotation and classified into four major density classes: chylomicrons, very low density lipoproteins (VLDL), low density lipoproteins (LDL) and high density lipoproteins (HDL) (Table 1.1).

The protein components of plasma lipoproteins are known as apolipoproteins, twelve of which have been identified (Alaupovic, 1971; Lee *et al.*, 1983). Apolipoproteins participate in lipoprotein synthesis, secretion, processing and catabolism. In addition to being important structural constituents of the lipoproteins, apolipoproteins function in the activation and inhibition of lipoprotein lipase (LPL), the activation of lecithin cholesteryl acyl transferase (LCAT), and in the binding of LDL particles to the low density lipoprotein receptor (for reviews see Zannis and Breslow, 1984; Assmann, 1985; see also Table 1.2).

The clinical and biological importance of the apolipoproteins has generated much interest in their structure, function and evolution. Apolipoprotein structure and function is summarized in Table 1.2, while details of chromosomal location and gene organization are presented in Table 1.3. In the following subsections individual

apolipoproteins will be discussed separately but Tables 1.2 and 1.3 should be referred to for detail.

1.2.1 Apolipoprotein A1

Apolipoprotein A1 comprises as much as 67% of the protein of HDL (Galton, 1985). APOA1 promotes cholesterol efflux from tissues and also acts as a cofactor for lecithin cholesterol acyl transferase (LCAT), which is responsible for virtually all plasma cholesterol esterification (Soutar *et al.*, 1975). APOA1 is synthesized in the liver and intestine as a 267 amino acid residue preproapoA1 (Zannis *et al.*, 1983). Cotranslational and post-translational cleavage of an 18 residue pre-peptide and a six residue pro-peptide results in a 243 residue mature apoA1 peptide. The gene for APOA1 clusters with APOC3 and APOA4 on chromosome 11. It is approximately 2.0kb in length and contains three introns (Law *et al.*, 1984; Karathanasis, 1985). The DNA sequence contains tandemly repeated 66bp regions (codons 99-230) which are thought to have arisen by gene duplication (Karathanasis *et al.*, 1983a). Exon 3 of the gene contains a block of 33bp which is homologous to a similarly positioned block of APOC2 (Luo *et al.*, 1986).

Table 1.1 Composition and properties of the human plasma lipoproteins. Modified after Zannis and Breslow, 1984, Table I, and Galton, 1985, Table 6.1.

Properties	Chylomicrons	VLDL	LDL	HDL
Density	<0.95	0.95-1.006	1.006-1.063	1.063-1.21
Electrophoretic mobility	origin	pre-beta	beta	alpha
Proteins(%)	1-2	6-10	18-22	45-55
Apolipoproteins† (% total protein)				
APOA1	7	minor	trace	67
APOA2	4	minor	trace	20
APOA4§	major	trace		minor
APOB	23	37	98	minor
APOC1	15	5-10	trace	1-3
APOC2	15	5-10	trace	1-3
APOC3	36	40	trace	3-5
APOD	present	minor	trace	1-5
APOE	minor	13	minor	present
Lipid composition (%)				
Triglyceride	85	55	7	5
Cholesterol	6	17	50	35
Phospholipid	9	28	43	60

† major > 5%, minor < 5%.
§ also present in lipoprotein free fraction of plasma.

Table 1.2 Properties of the major human apolipoproteins. Modified after Zannis and Breslow, 1984, Table III, and Galton, 1985, Table 6.2.

Apolipo-proteins	Molecular weight (kD)	Plasma conc.(mg/ml)	Isoelectric point	Site of synthesis	Function
APOA1	28.3	1.00-1.20	5.85-5.40	Intestine/liver	Activates LCAT
APOA2	17.0	0.30-0.50	5.00	Intestine/liver	Activates LPL Inhibits LCAT
APOA4	46.0	0.16	5.45	Intestine	Activates LCAT Modulates LPL
APOB-100	549.0	0.70-1.00	-	Liver	Binds to LDLR
APOB-48	246.0	-	-	Intestine	Triglyceride transport†
APOC1	6.3	0.04-0.06	7.5	Liver/intestine	Activates LCAT
APOC2	8.8	0.03-0.05	4.9	Liver/intestine	Activates LPL Inhibits LCAT
APOC3	8.8	0.12-0.14	4.7-5.0	Liver/intestine	Inhibits LPL, LCAT and hepatic lipase
APOD	33.0	0.06-0.07	5.0-5.2	Widespread	Activates LCAT
APOE	33.0	0.03-0.05	5.7-6.0	Widespread	Binds to LDLR and chylomicron remnant receptor Activates LCAT

† APOB is also essential for the assembly and secretion of VLDL and chylomicrons.

Table 1.3 Molecular properties of the major apolipoproteins. Modified after Luo *et al.*, 1986, Table 1.

Apolipo-proteins	Chromosomal location†	Length (codons) of coding region: Signal peptide	Pro-segment	Mature peptide	Total	Locations of introns§: Intron 1 (no.bp)	Intron 2 (no.bp)	Intron 3 (no.bp)	66bp repeats	Notes
A1	11q23-qter	18	6	243	267	-20 (197)	-4 (185)	+49 (588)	present	homology with APOC2
A2	1q21-q23	18	5	77	100	-24 (169-82)	-1 (293)	+44 (395)		dimer in human plasma
A4	11q23-qter				393				present	
B-100	2p24-p23					(357)	(77)	(none)		glycoprotein
B-48										glycoprotein
C1	19q12-q13.2	26	0	57	83	(138)	(1190)	(2630)	present	
C2	19q12-q13.2	22	0	79	101	-13 (2386)	-4 (167)	+50 (296)	present	homology with APOA1
C3	11q23-qter	20	0	79	99	-13 (600)	-2 (125)	+40 (1800)	present	glycoprotein
D	3p14.2-qter			169						glycoprotein homology with globulins
E	19q12-q13.2	18	0	299	317	-23 (700)	-4 (1100)	+61 (600)	present	glycoprotein

† Frezal and Klinger, 1987.
§ Intron 1, upstream from AUG preceding nucleotide number cited.
Intron 2, between first and second positions of codon number cited.
Intron 3, between second and third positions of codon number cited.
The first codon after the signal peptide cleavage site is taken as codon no. +1.

1.2.2 Apolipoprotein A2

Apolipoprotein A2 is the second most abundant component of HDL particles (20%), can displace APOA1 from HDL and has a role in the activation of hepatic lipase and in the inhibition of LCAT (Soutar *et al.*, 1975; Jahn *et al.*, 1983). APOA2 is located on chromosome 1 (Knott *et al.*, 1984*a*). The protein is synthesized in the liver and to a lesser extent the intestine as a 100 residue preproapoA2. The signal peptide consists of 18 amino acids and the propeptide of 5 amino acids (Tsao *et al.*, 1985). The mature protein is some 77 amino acids long and exists as a dimer in human plasma, linked by a disulphide bond at position 6 (Brewer *et al.*, 1972). The APOA2 gene contains three introns in positions similar to those of APOA1 (Breslow, 1988).

1.2.3 Apolipoprotein A4

Apolipoprotein A4 is a minor component of HDL and a major component of chylomicrons. It is also found in the lipoprotein free fraction of plasma. Indirect evidence points to a role for APOA4 in the biogenesis and/or secretion of intestinal triglyceride rich lipoproteins (Lackner *et al.*, 1985).

APOA4 is located on chromosome 11 some 12 kb 3′ to the APOA1 gene. Sequencing has revealed that APOA4 contains the 22mer periodicity already known to be present in APOA1. These repeats, common to a APOA1, APOA4 and APOE, give the proteins the potential to form amphipathic helices. It may be these structures in particular which are responsible for

the lipid binding potential of these apolipoproteins (Zannis and Breslow, 1984). The APOE*4 gene has three exons of 162, 127 and 1180 bp, but lacks an intron commonly found in other apolipoprotein genes in the 5′ untranslated region (Elshourbagy *et al.*, 1986, 1987).

1.2.4 Apolipoprotein B

In addition to being the principal constituent of LDL apolipoprotein B is also a major component of chylomicrons and VLDL. Virtually all of the 18-22% protein of LDL is apolipoprotein B (Herbert *et al.*, 1983). Apolipoprotein B is a glycoprotein with two forms, B-48 and B-100, derived from the intestine and liver respectively (Kane *et al.*, 1980). APOB-48 appears to correspond to the N-terminal 48% of APOB-100 (Olofsson *et al.*, 1987). Apolipoprotein B is the major protein determinant for the cellular recognition and catabolism of LDL by the LDL receptor. The specific lipid binding properties of APOB may be attributable to the presence of repeated amphipathic helical regions and of hydrophobic proline-rich domains (de Loof *et al.*, 1987). Apolipoprotein B is also essential for the assembly and secretion of VLDL and chylomicrons (Goldstein and Brown, 1983).

The gene for APOB maps to chromosome 2 (Knott *et al.*, 1986), is 43kb long and contains 28 introns (Blackhart *et al.*, 1986). The protein itself is some 4563 amino acids long and contains a 27 amino acid signal peptide (Knott *et al.*, 1986).

1.2.5 Apolipoprotein C1

Apolipoprotein C1 is a major constituent of VLDL, and is present in all other lipoprotein classes (Herbert *et al.*, 1983). The APOC1 protein acts as an activator for LCAT, but not as efficiently as apolipoprotein A1 (Soutar *et al.*, 1975). The 249bp translated portion of APOC1 is accounted for by a 57 amino acid mature peptide and a 26 amino acid prepeptide (Shulman *et al.*, 1975).

APOC1 is located within cluster of apolipoprotein genes on chromosome 19 along with APOC2, APOE and APOC1P1 (an APOC1 psuedogene). APOC1 has been localized 4.3kb 3′ to the APOE gene and contains three introns (Lusis *et al.*, 1986; Smit *et al.*, 1988).

1.2.6 Apolipoprotein C2

Apolipoprotein C2 is a component of VLDL and chylomicrons and has cofactor activity with lipoprotein lipase, the enzyme responsible for the hydrolysis of plasma triglycerides (Herbert *et al.*, 1983). The mature peptide of 79 amino acids is produced after cotranslational cleavage of a 22 amino acid signal peptide. The genomic structure of APOC2 has been determined (Wei *et al.*, 1985) and is similar to that of APOA1 with three introns in similar positions. As already mentioned APOC2 is a member of the chromosome 19 cluster and contains, within exon 3, a 33bp homology with APOA1.

1.2.7 Apolipoprotein C3

Apolipoprotein C3 is a major constituent of VLDL and is synthesized in the liver as a glycoprotein (Herbert *et al.*, 1983). It has been shown to inhibit both lipoprotein and hepatic lipase (Quarfordt *et al.*, 1982). Cotranslational cleavage of a 20 amino acid preapoC3 peptide yields a mature protein of 79 amino acids (Brewer *et al.*, 1974). The protein is present in the plasma as three isoforms, each of which contains a different amount of linked carbohydrate.

Apolipoprotein C3 is 3kb in length, and again shows three introns, a feature common to a number of the apolipoproteins. Intron 1 (600bp) is in the untranslated region, intron 2 (125bp) interrupts the prepeptide, and intron 3 (1.8kb) interrupts the mature protein (Karathanasis, 1985). The gene also carries internal repeats with structural resemblances to APOA2 (Luo *et al.*, 1986).

1.2.8 Apolipoprotein D

Apolipoprotein D, originally designated APOA3 (Kostner, 1974), is located on chromosome 3 and produces a glycoprotein as a constituent of HDL (Fielding and Fielding, 1980; Albers *et al.*, 1981). The 169bp mature protein shows no marked similarity with other apolipoproteins (hence the change in nomenclature) but rather has homology with genes of the α-globulin superfamily. Apolipoprotein D mRNA has been detected in human liver, intestine, spleen and fetal brain tissue - a pattern of synthesis dissimilar to most other

apolipoproteins, where synthesis is restricted to the liver and/or intestine (Drayna *et al.*, 1986).

The APOD gene is comprised of at least five exons, four of which have been sequenced. Exon 2 is 157bp long and encodes the 5′ untranslated sequence, the signal peptide and the codons for the first two amino acids. Exons 3 and 4 are 122 and 89bp long respectively. Exon 5 is 414bp in length and encodes nearly half of the mature protein plus all of the 3′ untranslated region of the mRNA (Drayna *et al.*, 1987*a*).

Most or all of LCAT is found complexed with APOD suggesting that APOD is active in the efflux of cholesterol from the peripheral tissues and the transport of cholesteryl esters to the liver for catabolism. Its association with lipids probably has a different basis than that found in the other apolipoproteins (Drayna *et al.*, 1986).

1.2.9 Apolipoprotein E

The major concentration of apolipoprotein E is found in the liver, adrenal gland and spleen, although it is synthesized in all major organs. Apolipoprotein E is a constituent of all lipoprotein classes and mediates lipoprotein catabolism extrahepatically via the LDL (or apoB/E) receptor and hepatically by LDLR and the chylomicron remnant receptor (Innerarity and Mahley, 1978). The three-dimensional structure of the apolipoprotein E protein has recently been determined and exhibits a four helix bundle, stabilized by a hydrophobic core, that includes leucine zipper type

interactions, and by numerous salt bridges (Wilson *et al.*, 1991).

The gene for APOE is located within the chromosome 19 cluster and produces a glycoprotein of 299 amino acids as a preapoE. PreapoE is cotranslationally cleaved of its 18 amino acid prepeptide and secreted into the plasma as a sialylated apolipoprotein (Zannis *et al.*, 1984). APOE consists of several isoproteins that differ in size and/or charge. This is the result of both genetic variation of APOE within the population and of post translational modification of APOE with carbohydrate chains containing sialic acid (Zannis and Breslow, 1981).

The gene is 3.7kb in length with three introns. Intron 1 (700bp) is in the 5′ untranslated region, intron 2 (1.1kb) interrupts the prepeptide, and intron 3 (600bp) interrupts the mature protein (Rall *et al.*, 1982). The pattern of 66bp tandem repeats is present (Das *et al.*, 1985) and striking homology exists between these repeats and those of human APOA1 and rat APOA4. The intron locations are very similar to those identified for APOA1, APOA2, APOC1, APOC2 and APOC3 and may indicate that each exon codes for a functionally distinct region of the protein. The similar splicing patterns (Fig 1.1) and homologies between many of the apolipoproteins have been cited as evidence in support of a common ancestral origin for the apolipoproteins (Barker and Dayhoff, 1977; Luo *et al.*, 1986). This super gene family does not include the APOB or APOD genes.

1.2.10 Other apolipoproteins

Apolipoprotein F has a molecular weight of 26,000 - 32,000 and an isoelectric point of 3.75. Along with APOG (MW 72,000), APOF is a minor component of HDL (Olofsson *et al.*, 1978). Apolipoprotein H or β2-glycoprotein (MW 43,000 - 50,000, 326 amino acids) is present in all lipoprotein classes (Polz and Kostner, 1979). APOH binds to chylomicrons and acts, with APOC2, in the activation of lipoprotein lipase (Nakaya *et al.*, 1980). Three common alleles for APOH have been identified in both Caucasoids and African Americans, with a fourth allele being observed at polymorphic frequencies in African American populations only (Kamboh *et al.*, 1988*a*).

1.2.11 Low density lipoprotein receptor

Low density lipoprotein receptors are high affinity receptors for LDL and as such internalize LDL by binding with apolipoproteins B and E (Brown and Goldstein, 1988). The LDLR also binds and internalizes β-migrating VLDL and intermediate density lipoproteins (IDL) (Falus and Romics, 1988). The receptor is present at some 50-100,000 copies per fibroblast cell, with 50-80% of the receptors being found clustered in coated pits on the cell surface.

The LDLR is a glycoprotein with a molecular weight of 160,000. It contains a signal peptide and five recognised domains and is coded for by a 45kb locus at 19p13.1-p13.3 (Lindgren *et al.*, 1985). The gene consists of 18 exons, the sequences of which bear strong correlations to the structural protein domains (Fig 1.2 and Sudhof *et al.*,

1985). The 3′ untranslated region has been shown to include several *Alu* repetitive elements (Yamamoto *et al.*, 1984; Hobbs *et al.*, 1985). As homology exists with epidermal growth factor (EGF), complement component (C9) and plasma proteases, the LDLR gene is thought to be a mosaic of coding sequences shared with other proteins. In all cases the shared sequences are encoded for by discrete exons, suggesting that LDLR is a member of at least two supergene families (Russell *et al.*, 1986).

1.2.12 Lipoprotein lipase

Lipoprotein lipase is responsible for the hydrolysis of core triglycerides in VLDL and chylomicrons, and as such provides fatty acids for storage in adipose tissue, or for oxidation in muscle and other tissues (Garfinkel and Schotz, 1987). The enzyme is found on the luminal surface of capillary endothelial cells and plays an important role in the maturation of lipoproteins. As a result alterations in its activity have an effect on both triglyceride and cholesterol levels.

The LPL gene is located at 8p22 and is a member of a gene family that includes hepatic lipase and pancreatic lipase (Wion *et al.*, 1987). The LPL gene codes for a mature protein of 448 amino acids (Wion *et al.*, 1987), and is approximately 30kb in length. It is comprised of ten exons, the first nine of which range in size from 105 to 276bp, with the tenth being 1948bp in length (Deeb and Peng, 1989; Kirchgessner *et al.*, 1989). Exon 1 codes for the signal peptide, exon 2 for the protein domain that is known to

bind to the lipoprotein substrate, and exons 6 and 9 are rich in basic amimo acids and are likely to be involved in anchoring the enzyme to the capillary endothelium (Deeb and Peng, 1989).

Figure 1.1 Apolipoprotein mRNA structure. Solid bar = protein coding regions; thin lines = lengths of 5′ and 3′ untranslated regions; solid arrowheads = location of introns. The beginning of all protein coding regions are aligned. From Breslow, 1988, Fig 6.

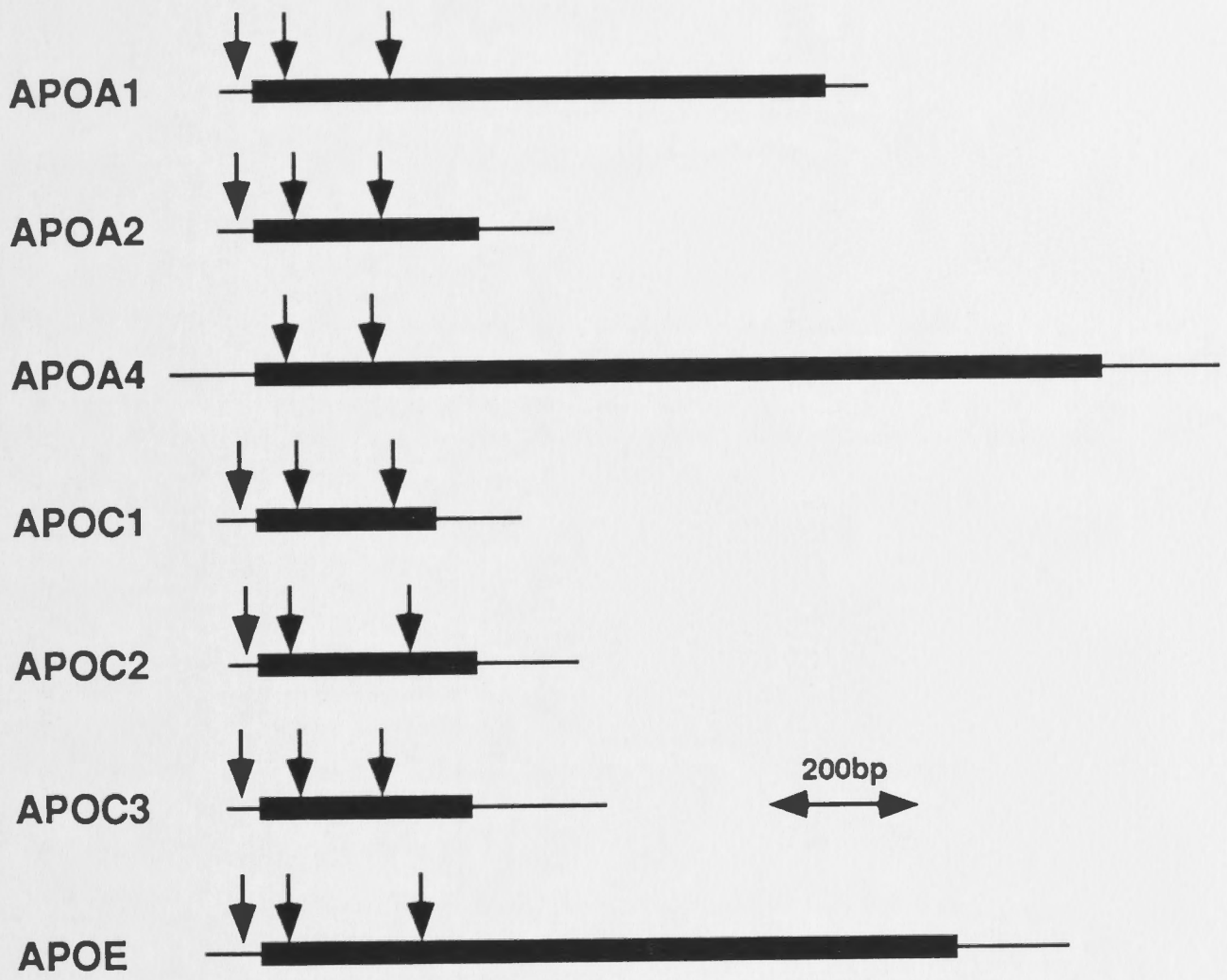
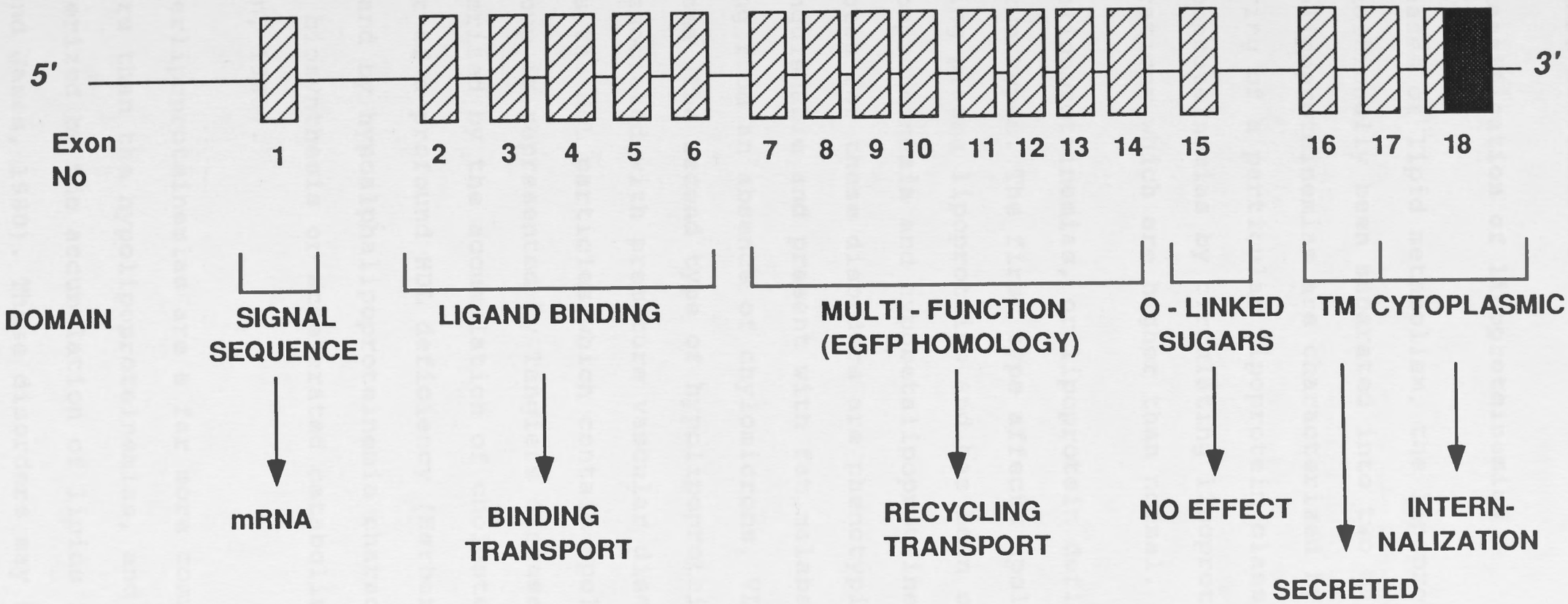

APOA1
APOA2
APOA4
APOC1
APOC2
APOC3
APOE
200bp

Figure 1.2 Low density lipoprotein receptor gene organization and exon structural domain relationship. Modified after Brown and Goldstein, 1988, Fig 3.


5'
3'
Exon No
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
DOMAIN
SIGNAL SEQUENCE
LIGAND BINDING
MULTI - FUNCTION (EGFP HOMOLOGY)
O - LINKED SUGARS
TM
CYTOPLASMIC
mRNA
BINDING TRANSPORT
RECYCLING TRANSPORT
NO EFFECT
SECRETED
INTERN-NALIZATION

1.3 DISEASE ASSOCIATIONS

1.3.1 Classification of lipoproteinemias

The diseases of lipid metabolism, the lipoproteinemias, have traditionally been separated into two major classes. The hypolipoproteinemias are characterized by the absence or lowering of a particular lipoprotein class, and the hyperlipoproteinemias by circulating lipoprotein concentrations which are higher than normal.

The hypolipoproteinemias, or lipoprotein deficiency states, are of two types. The first type affects apolipoprotein B containing plasma lipoproteins and has been subdivided into abetalipoproteinemia and hypobetalipoproteinemia. Homozygotes for these disorders are phenotypically indistinguishable and present with fat malabsorption resulting from an absence of chylomicrons, VLDL and LDL in the plasma. The second type of hypolipoproteinemia is more often associated with premature vascular disease and involves the HDL particles which contain apolipoprotein A. This group is represented by Tangiers disease which is characterized by the accumulation of cholesteryl esters as a result of a profound HDL deficiency (Herbert *et al.*, 1983), and by hypoalphalipoproteinemia characterized by reduced biosynthesis or accelerated catabolism of HDL (Assmann, 1985).

The hyperlipoproteinemias are a far more common group of disorders than the hypolipoproteinemias, and are characterized by the accumulation of lipids in the plasma (Gurr and James, 1980). These disorders may be monogenic or

polygenic in origin or may appear secondary to a systemic disease, a pattern of nutrition or lifestyle habit, such as excessive alcohol intake (Galton, 1985). Five major types of hyperlipoproteinemia are distinguished on the basis of the patterns of lipoprotein accumulation (Table 1.4). It must be stressed that each type is not a homogeneous entity from a clinical, genetic or pathophysiological point of view.

Table 1.4 Summary and classification of the hyperlipoproteinemias. Modified after Gurr and James, 1980, Table 6.2, Herbert *et al.*, 1983, Table 29-5, and Galton, 1985, Table 6.3.

Type	Lipoprotein changes	Raised lipids	Subtypes and/or primary causes	Clinical features† of primary disease	May be secondary to:
I	> chylomicrons < HDL, <LDL	> triglycerides	Lipoprotein lipase deficiency, APOC2 deficiency	heptosplenomegaly, eruptive xanthomas, lipaemia retinalis, pancreatitis	systemic lupus erythematosis, macroglobulinaemias
IIa	> LDL (normal)	> cholesterol	Familial hyper-cholesterolemia (LDLR defects) Familial polygenic hypercholesterol-emia	xanthomas, corneal arcus, premature vascular disease	excessive saturated fatty acid and cholesterol intake, hypothyroidism, nephrotic syndrome
IIb	> LDL (normal) > VLDL	> cholesterol > triglycerides	Familial combined hyperlipoprotein-emia, Lipoprotein lipase deficiency	as for IIa	as for IIa
III	> LDL (abnormal form=IDL§)	> cholesterol > triglycerides	defect in chylo-micron remnant removal,often associated with APOE*2/APOE*2	extensive vascular disease	diabetes, hypothyroidism, obesity

† Some or all features present, depending on subtype.
§ Intermediate density lipoprotein (IDL) particle present.
Continued overleaf.

Table 1.4 Cont'd. Summary and classification of the hyperlipoproteinemias. Modified after Gurr and James, 1980, Table 6.2, Herbert *et al.*, 1983, Table 29-5, and Galton, 1985, Table 6.3.

Type	Lipoprotein changes	Raised lipids	Subtypes and/or primary causes	Clinical features† of primary disease	May be secondary to:
IV	> VLDL	> triglycerides	Familial combined hyperlipoproteinemia, Familial hypertriglyceridemia, Polygenic hyperlipoproteinemia Lipoprotein lipase deficiency	vascular disease (expression greatly affected by environmental factors e.g. obesity)	diabetes§ exogenous oestrogen therapy§ excessive alcohol intake§ renal disease hypothyroidism
V	> chylomicrons > VLDL	>triglycerides (higher than IV)	Familial hypertriglyceridemia, sometimes associated with APOE*4, Lipoprotein lipase deficiency	vascular disease, heptosplenomegaly xanthomas, lipaemia retinalis, pancreatitis (75% with impaired glucose tolerance)	as for type IV, secondary to many diseases and environmental factors

† Some or all features present, depending on subtype.
§ Often with underlying lipid intolerance.

1.3.2 Lipoprotein gene variation and associations with lipoproteinemias

No variants or associations with the lipoproteinemias are documented for APOA2, APOC1 or APOD. A number of variants and associations have been found for the remaining apolipoproteins and for the LDLR and LPL. These will be discussed in the following section.

1.3.2.1 Apolipoprotein A1

The estimated prevalence of apolipoprotein A1 mutations is 1:1,000,000. Some variants are associated with severe hypolipoproteinemias, whilst others alter APOA1 levels but do not predispose to heart disease (Assmann, 1985). Two variants, APOA1*Milana and APOA1*Marburg, are associated with decreases in HDL levels and consequent hypoalphalipoproteinemia, and three, APOA1*Geissen, APOA1*Marburg and APOA1*Munster 2a are defective in the activation of LCAT. The remainder of the APOA1*Munster variants (1, 3 and 4) lower serum levels of APOA1 by 50% but seemingly have no other effect (Breslow, 1985).

The recent identification of nine new variants and the structural analysis of the now twenty known variants has revealed that proline to arginine, and proline to histidine substitutions are over-represented, and that hypervariablilty of the apolipoprotein A1 α-helical domain may result from the occurrence of CpG dinucleotides in this region (von Eckardstein *et al.*, 1990).

Tangiers disease plasma contains very low levels of apolipoprotein A1. However, many studies have failed to

find any APOA1 structural mutation which could be responsible for the disease. There has, though, been one report of an altered apolipoprotein A1 isoform being found in Tangiers patients (Bisgaier *et al.*, 1987).

Complete APOA1 (or APOA1-C3) deficiency results in the absence of HDL, presence of multiple xanthomas and early onset of CHD. In a well documented case the deficiency is due to an insertion of a 6.5kb segment (which contains some APOC3 DNA) into the 3′ end of the 3rd intron of APOA1 (Karathanasis *et al.*, 1983*b*).

1.3.2.2 Apolipoprotein A4

Seven protein variants, detectable by isoelectric focussing, have been described (Kamboh and Ferrell, 1987; Sepehrnia *et al.*, 1988*a*). Neither heterozygosity nor homozygosity for minor alleles (APOA4*2 to APOA4*7) show association with plasma lipoprotein abnormalities or atherosclerosis susceptibility (Breslow, 1985).

1.3.2.3 Apolipoprotein B

An unknown recessive defect of APOB has been identified which leads to an absence of APOB from HDL and the occurrence of either abeta- or hypobeta-lipoproteinemia (Assmann, 1985). In contrast the over-production of APOB may be involved in familial combined hyperlipoproteinemia (Janus, *et al.*, 1980). Some individuals are known to produce chylomicrons but have low to absent LDL-cholesterol. These individuals have been shown to produce

intestinal APOB-48 but not hepatic APOB-100 (Malloy, *et al.*, 1981).

A series of five allelic variants called Ag have been defined which are closely linked to APOB. One antigenic determinant Ag(x) is associated with variation in triglyceride and cholesterol levels, but not with an increase in atherosclerosis risk (Berg *et al.*, 1986). The Lp(a) protein, originally thought to be an antigenic determinant, has now been shown to be a distinct protein which is bound to APOB by disulphide bonds. Increased levels of Lp(a) have been associated with CHD and cerebrovascular disease (Murai *et al.*, 1986).

1.3.2.4 Apolipoprotein C2

A deficiency of APOC2 is present in a small proportion of individuals with type I hyperlipoproteinemia. These individuals may initially be diagnosed as having lipoprotein lipase deficiency, but it is the absence of APOC2, a cofactor for LPL, that is the cause of the disorder in these cases (Yamamura, *et al.*, 1979). The molecular basis in two of the deficient APOC2 alleles has been determined. APOC2*Nijmegen results from a single base deletion of G at position 2943, causing a frame shift and introducing a premature termination codon (Fojo *et al.*, 1988*a*). A G to C substitution within the donor splice site of intron 2 has been identified in the APOC2*Hamburg allele (Fojo *et al.*, 1988*b*). Three protein variants of unknown effect have also been described (Sepehrnia *et al.*, 1988*b*).

1.3.2.5 Apolipoprotein C3

The three isoforms of APOC3 are the result of varying amounts of attached neuramic acid. APOC3*0, APOC3*1 and APOC3*2 contain zero, one and two attached residues of neuramic acid respectively. Normal individuals have a preponderance of APOC3*0. Type V hyperlipoproteinemics have a lowering in the proportion of APOC3*0, whilst some cases of abetalipoproteinemia and one type of hypobetalipoproteinemia have been shown to lack the APOC3*1 isoprotein altogether (Herbert *et al.*, 1983). The complete absence of APOC3 is found very rarely, as in the combined APOA1-C3 deficiency already mentioned.

1.3.2.6 Apolipoprotein E

Apolipoprotein E has three major alleles that contribute to the normal variance of plasma lipoproteins and which are determined by the amino acids at residues 112 and 158. The APOE*2 allele has a cysteine at both positions, the APOE*3 allele has a cysteine at position 112 and an arginine at position 158, and APOE*4 has an arginine at both positions.

This APOE polymorphism has a pronounced effect on LDL cholesterol, total plasma cholesterol, APOB and APOE levels. Moreover, this effect appears to be independent of ethnicity. Low serum cholesterol levels and LDL cholesterol levels are detected in APOE*2 homozygotes, and high levels in APOE*4 homozygotes (Utermann, 1985). The APOE*2 allele is also associated with high levels of plasma triglycerides (Assmann *et al.*, 1984; Sing and Davignon, 1985; Ehnholm *et al.*, 1986). The APOE polymorphism accounts for 20% of the

variability of APOE levels, 12% of the variability of APOB levels and 4% of the variability of total-cholesterol (Boerwinkle and Utermann, 1988).

In Type V hyperlipoproteinemia 74% of patients possess at least one APOE*4 allele, while in Type III 90% of sufferers are APOE*2/APOE*2 homozygotes (Utermann *et al.*, 1979; Breslow *et al.*, 1982a). However only 1-2% of people with the APOE*2/APOE*2 genotype go on to express the disease. Type III hyperlipoproteinemia becomes manifest only when, in addition to APOE*2 homozygosity, there is a further inborn error of lipid metabolism present or a secondary hypolipoproteinemia (Assmann, 1985). The APOE*2 allele is significantly more frequent in hypertriglyceridemia, the APOE*4 in hypercholesterolemia. However, both alleles are found more frequently in mixed hyperlipoproteinemia than in normal individuals, with 20% of patients having at least one rare allele (Utermann *et al.*, 1984).

1.3.2.7 Low density lipoprotein receptor

Familial hypercholesterolemia (FH) is characterized by the impaired removal of LDL from blood and results from mutations in the low density lipoprotein receptor. Heterozygotes occur at a frequency of 1 in 500 in most populations, have LDL levels on average two times that of normal individuals and are likely to suffer from CHD by the age of fifty. Homozygotes have LDL levels up to six times the normal levels and develop CHD in the first or second decade of life (Brown and Goldstein, 1988).

The clinical symptoms are the result of four classes of mutations of the receptor, each of which has a different phenotypic effect on the protein:-

1) Null alleles. One third of FH homozygotes show no detectable LDLR protein and only minimal amounts of LDLR mRNA.

2) Transport deficient alleles. The improper folding of the protein results in the accumulation of LDLR in the endoplasmic reticulum.

3) Binding deficient alleles. LDLR reaches cell surface but fails to bind LDL.

4) Internalization defective alleles. A defect in the cytoplasmic domain prevents targeting of the LDLR to coated pits on the cell surface.

Sixteen individual mutations had been mapped or sequenced and assigned to the four classes by 1986 (Russell *et al.*, 1986). Although some populations, such as the Afrikaners in South Africa and the French Canadians, carry only a few LDLR mutations, the mutational heterogeneity at the locus has made it difficult to apply DNA techniques for use in population screening (Motulsky, 1989). This problem is being addressed with the application of the polymerase chain reaction and the examination of amplified fragment length polymorphisms at the locus (Savolainen *et al.*, 1991).

1.3.2.8 Lipoprotein lipase

The genetic lesions responsible for a subset of type I hyperlipoproteinemia (lipoprotein lipase deficiency) have been determined in a number of families. These lesions represent a range of molecular events including gene duplication (Devlin *et al.*, 1990), point substitution (Emi *et al.*, 1990*a*, 1990*b*; Hata *et al.*, 1990*a*) and insertions (Langlois *et al.*, 1989).

1.3.3 Restriction fragment length polymorphism associations of the lipoprotein genes with hyperlipoproteinemia and cardiovascular disease

From the previous section it is apparent that several of the hyperlipoproteinemias are associated with quantitative and/or qualitative alterations of the lipoprotein genes. The genetic mechanisms involved in many of the subtypes of hyperlipoproteinemia remain unclear. Associations with the hyperlipoproteinemias, and more generally with atherosclerosis, have been sought with RFLPs identified by the cloned lipoprotein genes. A selected review of this work is presented in Table 1.5.

Most interest has focussed on the APOA1-C3-A4 cluster of chromosome 11. At least 11 RFLPs occur at this locus and, of these, six have been used in clinical studies (Antonarakis *et al.*, 1986). Rees *et al.* (1985) claimed a high frequency (relative incidence=57) of the rare allele (S2) for the polymorphic site of *Sst*I in the APOC3 gene in hypertriglyceridemics (Types IV and V) when compared with controls in Caucasoids from an English population. This association has not been found in either Type IIb (Rees *et*

al., 1983) or III (Vella *et al.*, 1985). A large difference in S2 allele frequency has also been seen between survivors of myocardial infarct (MI) and controls (Ferns *et al.*, 1985). However the association appears to be population specific since no associations have been seen in Scottish (Ferns *et al.*, 1986a; Price *et al.*, 1986), Norwegian (Kessling *et al.*, 1986), Chinese, African or Indian/Asian populations (Rees *et al.*, 1985).

Table 1.5. Associations of lipoprotein abnormalities and atherosclerosis with apolipoprotein gene RFLPs. From Breslow, 1988, Table 7.

Apolipo-protein	Enzyme	Associated phenotype
APOA1	*Xmn*I	MI resistance, CAD resistance, high APOA1 levels, HLP types IIb,III, V, high TG levels
	*Msp*I	MI, CAD resistance
	*Pst*I	Premature CAD, MI (<60yrs), MI resistance (>60yrs), low HDL levels
	*Eco*RI, *Pst*I	CAD, low HDL levels
	*Bam*HI	
	*Hind*III#	
APOC3 (APOA1)	*Sac*I/*Sst*I	MI, CAD, HLP type V, low HDL levels, high TG levels
	*Pvu*II	CAD
APOA2	*Msp*I	High APOA2 levels, high TG levels
APOB	*Xba*I	MI, low TG and Chol levels, Ag(x), Ag(c)/Ag(g) locus
	*Eco*RI	MI, Ag(t)/Ag(z) locus
APOC2	*Taq*I	HLP type I (APOC2 deficiency)
APOE	*Hpa*I	HLP type III

RFLP, restriction fragment length polymorphous; HDL, high density lipoprotein; HLP hyperlipoproteinemia; MI, myocardial infarction; CAD, coronary artery disease; TG, triglyceride, Chol, cholesterol; Ag(c), Ag(g), Ag(t), Ag(x) and Ag(z) are determinants of the Ag allelic system.
Detects insertion/deletion.

In an attempt to clarify the situation additional RFLPs around the APOA1-C3-A4 cluster have been investigated to enable haplotype analysis. Whilst a *Pst*I RFLP showed no

association with hyperlipoproteinemia (Kessling *et al.*, 1985; Ferns and Galton, 1986*a*; Hayden *et al.*, 1987) an association has been claimed with lowered HDL levels and premature coronary artery disease (Anderson *et al.*, 1986; Ferns and Galton, 1986*a*). *Xmn*I (Kessling *et al.*, 1985; Hayden *et al.*, 1987) and *Msp*I (Ferns and Galton, 1986*a*) RFLPs in this complex show only weak association with hyperlipoproteinemics and MI survivors. A specific haplotype of *Sst*I and *Msp*I, S1-M2 was significantly increased in Caucasoid and Japanese MI survivors (Stocks *et al.*, 1987). A rare *Eco*RI allele has been reported to be more frequent in patients with clinical symptoms of CHD than normal controls (Buraczynska *et al.*, 1986). A 300bp deletion of an *Alu* element 5kb upstream to the APOA1 gene occurs with a frequency of 0.05 and 0.20 in Germans and North American Caucasoids respectively (Frossard *et al.*, 1986). This deletion is detected by a number of RFLPs and is associated with angiographically proven atherosclerosis and decreased HDL cholesterol levels (Lim *et al.*, 1986).

One association between low HDL levels and increased triglyceride levels and an *Msp*I RFLP of APOA2 has been reported (Deeb *et al.*, 1986). This report supports an earlier finding of altered HDL composition in homozygotes for the possession of the site (Scott *et al.*, 1985).

Presence of the APOB *Xba*I restriction site is reported to be associated with a 36% increase of triglycerides in normal male Caucasoids (Deeb *et al.*, 1986; Law *et al.*, 1986), but this has not been verified in a study of MI

survivors (Ferns and Galton, 1986*b*). Homozygotes for the absence of this site have lower levels of APOB, total serum cholesterol, triglycerides and LDL-cholesterol (Leren *et al.*, 1988). Hegele *et al.* (1986) examined the *Eco*RI and *Msp*I RFLPs in addition to *Xba*I and found a significant increase of the presence of the sites in Caucasoid MI survivors. This finding was not repeated for *Xba*I in Japanese (Aburatani *et al.*, 1987). Finally, Deeb *et al.* (1986) have reported associations between *Taq*I and *Pvu*II sites, increased cholesterol levels, and decreased HDL levels.

A *Taq*I RFLP of APOC2 appears associated with increased cholesterol levels (Deeb *et al.*, 1986) although in a review (Breslow, 1988) it is claimed that neither the *Taq*I RFLP, nor a closely linked *Bgl*I RFLP, show any association with clinical abnormalities.

The presence of the *Pvu*II site in LDLR has been found to be associated with decreased cholesterol, LDL-cholesterol and triglyceride levels (Brink *et al.*, 1986; Humphries *et al.*, 1991). Homozygosity for the presence of a *Pvu*II site in LPL has been found to be associated with decreased triglyceride levels (Chamberlain *et al.*, 1989).

Comparisons between studies of this type are often hampered by the number of alternate classifications for the disease states and the method of selecting the "control" individuals. These may be randomly chosen from the general population or actively chosen to exclude the disease state. Many minor allele frequencies differ greatly among races;

controls and patients must be racially matched. In addition, the delineation of plasma lipid concentrations into normal and hyperlipoproteimemic classes is arbitrary. It may be more appropriate to determine the average allelic effect on continuous phenotypic variation in plasma lipoprotein levels (Boerwinkle and Utermann, 1988) than to attempt to establish allelic associations with heterogeneous clinical subtypes.

The above listed difficulties in the planning and interpretation of this type of study have led to many contradictory and non-repeatable associations being presented in the literature. For example associations cited in Table 1.5 (from Breslow, 1988) are in part contradicted in other reviews (Deeb *et al.*, 1986; Cooper and Clayton, 1988). A great deal more work is required before there is agreement on the extent and effects of RFLP disease associations in this field.

1.3.4 Genetics of non-insulin dependent diabetes mellitus

Concordance for NIDDM in monozygotic twins approaches 100% (Barnett *et al.*, 1981; Kuzuya *et al.*, 1988). Family clustering of NIDDM is evident with more than 30% of the siblings of individuals with NIDDM having abnormal glucose tolerance (Kobberling and Tillil, 1982; O'Rahilly *et al.*, 1987). It is therefore accepted that NIDDM has a strong genetic component.

Indirect evidence for genetic component in NIDDM is provided by admixture studies in North Dakota Amerindians,

Nauruans, and Mexican Americans (Brosseau *et al.*, 1979; Serjeantson *et al.*, 1983; Chakraborty *et al.*, 1986). These studies reveal a decrease in prevalence of NIDDM with increasing foreign admixture. Bimodality of glucose tolerance distributions among Nauruans and Pima Indians has also been interpreted as indirect evidence for the presence of a genetic component in the aetiology of NIDDM (Rushforth *et al.*, 1971; Zimmet and Whitehouse, 1978).

A major gene has been implicated in NIDDM in the Nauruan population. Only 5% of 20 offspring from non-diabetic parents where hyperglycaemic, compared with 79% of 43 offspring from diabetic parents (Serjeantson and Zimmet, 1984). Segregation analysis of age adjusted fasting glucose and post-load glucose levels favours a dominant autosomal mode of inheritance but is also compatible with a codominant mode of inheritance (Serjeantson and Zimmet, 1989).

Candidate genes that have been examined in relation to NIDDM include the insulin gene, the insulin receptor gene, the family of glucose transporter genes, lipoprotein genes, and the amylin gene. Possible associations between blood group markers, HLA, and serum protein markers and NIDDM have also been investigated (Serjeantson and Zimmet, 1991). The adenosine deaminase locus, recently reported to be in linkage with MODY (maturity-onset diabetes of the young) (Bell *et al.*, 1991) has not been implicated in NIDDM. Many studies have reported associations with NIDDM, with subsequent reports refuting these findings. There are as

yet no established RFLP markers for NIDDM (Serjeantson and Zimmet, 1991).

1.3.5 Altered lipoprotein metabolism in non-insulin dependent diabetes mellitus

Non-insulin dependent diabetes is accompanied by alterations in lipoprotein metabolism (Brunzell *et al.*, 1985; Howard, 1987; Assmann and Schulte, 1988). LDL particles contain approximately 70% of total plasma cholesterol and tend to be raised in poorly controlled diabetic patients (Betteridge, 1989). As insulin increases the number of LDL receptors (Mazzone *et al.*, 1984), chronic insulin deficiency might be associated with a diminished number of LDL receptors which would result in a decrease in the clearance of LDL. In addition, glucosylation of lysine residues on the LDL apolipoprotein B occurs in diabetic patients at three to four times the rate seen in normal individuals (Curtiss and Witztum, 1983, 1985). It has not been established whether this amount of glucosylation affects LDL binding (Steinbrecher and Witztum, 1984; Schleicher *et al.*, 1985). NIDDM LDL has been found to be enriched with triglycerides (Schonfeld *et al.*, 1974) and LDL from patients with NIDDM and hypertriglyceridemia has shown reduced binding to cultured skin fibroblasts (Hiramatsu *et al.*, 1985).

High density lipoprotein metabolism is also altered in NIDDM. HDL cholesterol concentrations tend to be low in NIDDM patients (Nikkila, 1978), and there is a stronger association of HDL with vascular disease in NIDDM patients

(Reckless *et al.*, 1978). These alterations may be the result of a decrease in the rate transfer of surface triglycerides to HDL, brought about by the lowering of lipoprotein lipase activity in the presence of insulin deficiency (Nikkila and Hormila, 1978). Other factors must also influence HDL metabolism in NIDDM as HDL levels often remain low even after hypertriglyceridemia and hyperglycaemia have been corrected (Betteridge, 1989).

The production of VLDL is increased in NIDDM (Abrams *et al.*, 1982). It is possible that hyperinsulinemia associated with insulin resistance stimulates VLDL production. Alternatively, it is proposed that it is insulin deficiency at the cellular level that leads to the over-production of VLDL (Gibbons, 1986).

Changes in lipid and lipoprotein levels in NIDDM patients can be moderated by improving glycaemic control. However, in patients with moderate to poor glycaemic control, and associated obesity, HDL levels will often be low and triglyceride, and possibly cholesterol levels, will be high (Betteridge, 1989).

1.3.6 Restriction fragment length polymorphism associations of the lipoprotein genes with non-insulin dependent diabetes mellitus

A relatively small number of studies have been undertaken examining lipoprotein RFLPs in NIDDM. The *Sst*I RFLP of APOA1-C3 revealed no differences in frequency between NIDDM and normal Caucasoids (Jowett et al., 1984). In contrast, Trembath (1987) has found an association between the S2

allele in NIDDM subjects with concurrent cardiac heart disease versus normals (no ethnicity cited). An earlier study by Buraczynska *et al.* (1985), of a 2.5kb *Eco*RI fragment in the APOA1 region, revealed no association with atherosclerosis, but did show an association with NIDDM.

Xiang *et al.* (1989) have investigated RFLPs at the APOA2, APOB and APOA1/C3/A4 locus and report a significant association with NIDDM of the APOB *Xba*I RFLP in individuals of lean/normal weight. Significant differences in the distribution of *Msp*I/*Pst*I/*Sst*I haplotypes at the APOA1/C3/A4 locus are also reported between overweight diabetic and non-diabetic groups. Most of this difference is observed with the *Msp*I RFLP.

An increased frequency of APOE*2 homozygotes among NIDDM patients has been observed by Vogelberg and Maucy (1988). The APOE*2 and APOE*4 alleles have also been reported to be more frequent in diabetics with hyperlipoproteinemia (Parhofer *et al.*, 1990).

1.4 STUDY POPULATION - THE WESTERN PACIFIC REPUBLIC OF NAURU

The island of the Republic of Nauru is situated in the Central Pacific just south of the equator, at longitude 160°55′ east and latitude 0°32′. The island is approximately 20 miles in circumference with an area of 6263 acres and is positioned at the intersection of the three Pacific ethnic regions of Melanesia, Micronesia and Polynesia (Figure 1.3) (Zimmet *et al.*, 1977). There is an indigenous population of Micronesian ancestry of

approximately 5500 persons. Extensive mining of phosphate deposits on the island has resulted in Nauru having one of the highest per capita incomes in the world and the lifestyle of the Nauruans has consequently changed, resulting in obesity and reduced physical activity and a diet of imported processed Western food, (Zimmet *et al.*, 1990).

Diabetes prevalence was found to be extremely high in the adult Nauruan population; in a population based study the prevalence in over 20 year olds was 30.3% (Zimmet *et al.*, 1977), some 13 times higher than in the Australian population (Welborn *et al.*, 1968). This high prevalence of diabetes, which was exclusively of the non-insulin dependent type, has only been exceeded in the American Pima Indians (Knowler *et al.*, 1978). Diabetic prevalence was rare in the first half of the century, with mortality rates due to diabetes showing an exponential increase since the 1950s (Schooneveldt *et al.*, 1988). This high prevalence of NIDDM has been confirmed in a larger study and the annual incidence estimated to be 1.6% (Zimmet *et al.*, 1984; Balkau *et al.*, 1985). Age standardized prevalence of NIDDM has remained constant (24.0% in 1987), but the incidence of impaired glucose tolerance has decreased from 21.1% in 1975/1976 to 8.7% in 1987 (Dowse *et al.*, 1991).

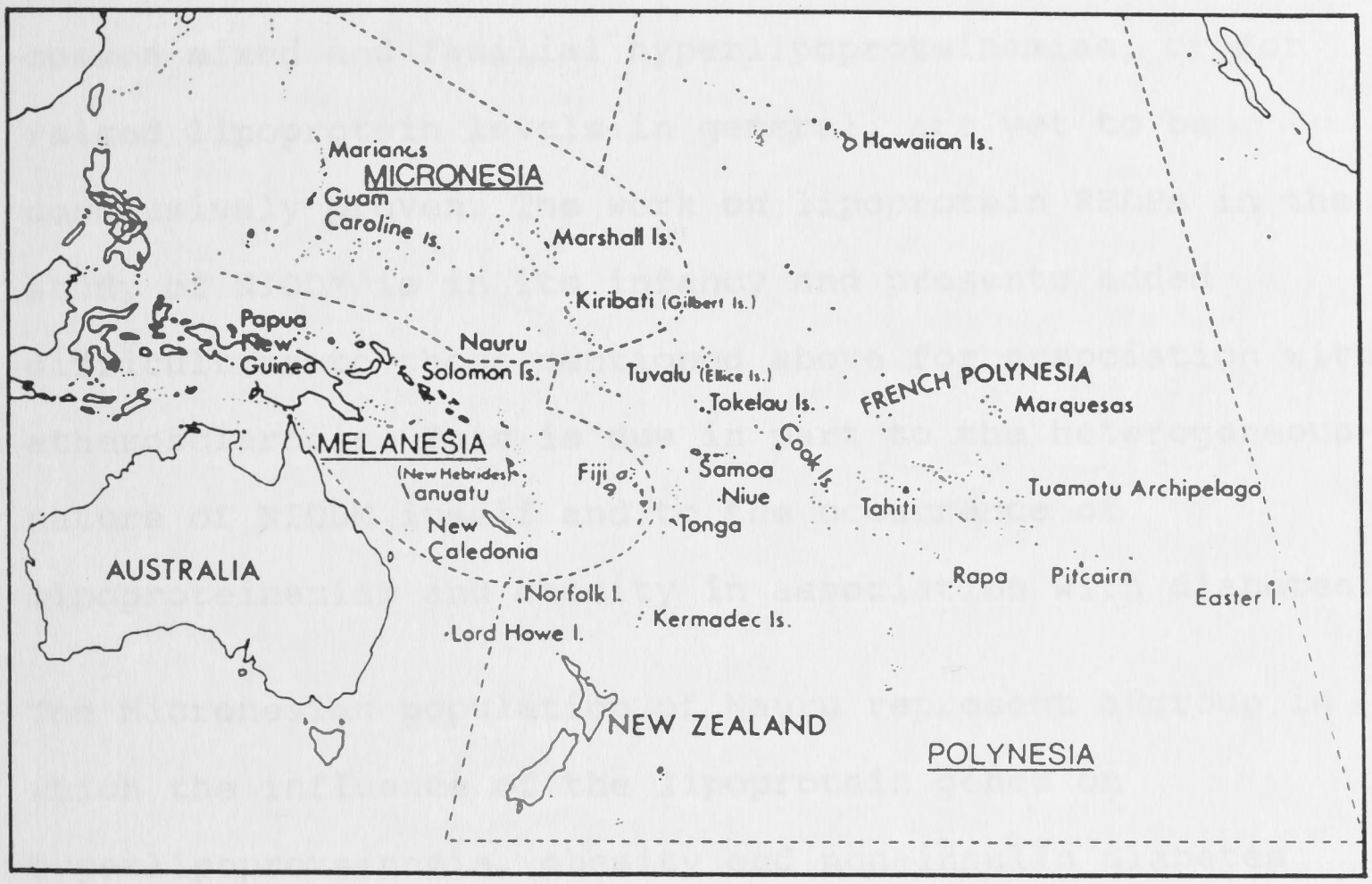
An analysis of the risk factors for NIDDM in this population has been undertaken (Balkau *et al.*, 1985). The 2hr post glucose load plasma glucose level was found to be the strongest determinant of diabetes in both sexes. In

females obesity, 2hr plasma insulin, fasting plasma glucose and uric acid were also found to be predictors for diabetes. Among males the predictors were age and plasma triglyceride levels, with obesity having marginal significance.

There is evidence for a heightened susceptibility to NIDDM in the Nauruan population. Foreign admixture lowers the occurrence of NIDDM (Serjeantson *et al.*, 1983). Bimodality exists in plasma glucose distributions (Zimmet and Whitehouse, 1978) and the prevalence of NIDDM in Nauru, after adjustment for age and obesity, is three times higher than that seen in a neighbouring Micronesian population (King *et al.*, 1984).

In addition the population exhibits a high rate of obesity, which has long been considered a risk factor for NIDDM (WHO Expert Committee, 1980). Body mass index has been shown to be a predictor of NIDDM in Nauruans and Japanese (Kadowaki *et al.*, 1984; Sicree *et al.*, 1987) although this is not the case for the Pima Indians (Saad *et al.*, 1988). The influence of obesity on the occurrence of NIDDM remains undefined, and it is now suggested that the duration of obesity, and the degree of central adiposity, may be more important factors in determining the predisposition of an individual to diabetes (Modan *et al.*, 1986; Zimmet *et al.*, 1990).

Figure 1.3 Map of the Pacific region illustrating major geographical and ethnic divisions. The island of Nauru is situated close to the intersection of Micronesia, Melanesia and Polynesia. From Zimmet *et al.*, 1990 , Figure 1.


Marianas
MICRONESIA
Guam
Caroline Is.
Marshall Is.
Hawaiian Is.
Kiribati (Gilbert Is.)
Papua New Guined
Nauru
Solomon Is.
Tuvalu (Ellice Is.)
Tokelau Is.
FRENCH POLYNESIA
Marquesas
MELANESIA
(New Hebrides)
Vanuatu
Fiji
Samoa
Niue
Cook Is.
Tuamotu Archipelago
Tahiti
Tonga
New Caledonia
AUSTRALIA
Rapa
Pitcairn
Easter I.
Norfolk I.
Kermadec Is.
Lord Howe I.
NEW ZEALAND
POLYNESIA

1.5 CONCLUSIONS

Whilst many studies have investigated the role of the apolipoproteins, lipoprotein lipase and the low density lipoprotein receptor with respect to rare lipid disorders, sometimes producing definite correlations with disease states, many of the associations claimed for the more common mixed and familial hyperlipoproteinemias, or for raised lipoprotein levels in general, are yet to be conclusively proven. The work on lipoprotein RFLPs in the study of NIDDM is in its infancy and presents added difficulties to those mentioned above for association with atherosclerosis. This is due in part to the heterogeneous nature of NIDDM itself and to the occurrence of lipoproteinemias and obesity in association with diabetes.

The Micronesian population of Nauru represent a group in which the influence of the lipoprotein genes on hyperlipoproteinemia, obesity and non-insulin diabetes mellitus can be examined. A high prevalence of NIDDM and obesity occurring against a relatively homogeneous ethnic and environmental background provides an opportunity to investigate the relative contribution of the lipoprotein genes to these multifactorial disorders.

1.6 AIMS

The aim of this study is to explore the influence of lipoprotein gene variation on non-insulin dependent diabetes mellitus, hyperlipoproteinemia and obesity in the Micronesian population of Nauru. The degree of variability of a number of lipoprotein gene fragment length

polymorphisms will be established and compared with that seen in other ethnic groups. The suitability of these genes as candidate genes for non-insulin dependent diabetes mellitus and hyperlipoproteinemia will be investigated. The effect of the variability present at these loci on diabetic status, plasma cholesterol and plasma triglyceride levels, and on measures of obesity, will be determined in both univariate and multivariate analyses. Linkage studies will be undertaken to examine the possibility of linkage between the lipoprotein loci and non-insulin dependent diabetes mellitus or hyperlipoproteinemia in the Micronesian population.

SECTION 2

MATERIALS AND METHODS

2.1 THE WESTERN PACIFIC REPUBLIC OF NAURU 1982 AND 1987 SURVEYS

In January of 1982 83% (n=1583) of adult Micronesians over the age 20yrs took part in a comprehensive baseline cardiovascular and diabetic survey on the Western Pacific Republic of Nauru (Zimmet *et al.*, 1984). In February 1987 all subjects were invited to participate in a follow-up survey and a sample of n=1184 (72.7% of the eligible sample) was obtained (Collins *et al.*, 1989). These surveys were undertaken by the World Health Organization Collaborating Centre for the Epidemiology of Diabetes Mellitus, Royal Southern Memorial Hospital, Melbourne. Survey data were collected with the co-operation and assistance of the Department of Health, Nauru, and the people of Nauru. Blood samples and survey data were made available to the Human Genetics Group, John Curtin School of Medical Research, Canberra, as part of a collaborative research project with the World Health Organization Collaborating Centre for the Epidemiology of Diabetes Mellitus.

2.1.1 Pedigree data

Participants in the 1982 survey provided the names of their parents allowing pedigrees for the populations to be constructed. Prospective pedigrees were subsequently checked by Nauruan informants. All two parent pedigrees were checked for veracity with genetic markers (Serjeantson and Zimmet, 1984).

2.1.2 Diabetic status

Plasma glucose concentrations were measured in both surveys following an overnight fast and again two hours after an oral 75g glucose load. Plasma glucose levels were measured on site with a glucose analyser (YSI, Yellow Springs, OH, USA) which uses a glucose oxidative method. Diabetic status was ascertained based on World Health Organization (WHO) recommendations (WHO Expert Committee, 1980; WHO Study Group, 1982) and following the criteria of the National Diabetes Data Group (1979). All subjects with a previous classification of diabetes mellitus were classified as such. A 2hr plasma glucose reading of ≥11.1 mM, or if no 2hr reading was available, a fasting glucose level of ≥7.8mM led to a diagnosis of diabetes mellitus. Impaired glucose tolerance was indicated by a 2hr glucose level of ≥7.8 mM and <11.1 mM. Subjects with fasting and 2hr glucose levels below 7.8mM were classified as having normal glucose tolerance.

2.1.3 Morphometric data and lipid levels

Body mass index was calculated in both surveys as weight over height squared (kg/m^2). Waist and hip measurements were collected at the 1987 survey only and a waist/hip ratio was calculated for each individual. Fasting plasma cholesterol and triglyceride levels were obtained in both surveys using a SMAC autoanalyser (Technicon, Tarrytown, NY, USA) (Zimmet *et al.*, 1984; Collins *et al.*, 1989).

2.2 REAGENTS AND LABORATORY MATERIALS

2.2.1 Reagents

The following list itemizes reagents of particular importance and the associated company of supply:-

[α-^{32}P]dATP (3000 mCi/mM); Amersham, Amersham, UK
acrylamide; Sigma Chemical Co., St Louis, MO, USA
agar; Difco Laboratories, Detroit, MI, USA
agarose; Type II, Sigma
ampicillin; Sigma
bacto-tryptone; Difco Laboratories
bacto-yeast extract; Difco Laboratories
black and white instant film (T55, T57); Polaroid Co., Cambridge, MA, USA
chloramphenicol; Sigma
dextran sulfate; Pharmacia, Uppsala, Sweden
ethidium bromide; Sigma
Nick translation kit (Code N.5500); Amersham
N,N'-methylene-*bis*-acrylamide; Bio-Rad Laboratories, Richmond, CA, USA
N,N,N',N' tetramethylethylene diamine (TEMED); Sigma
nylon hybridization membranes:- GeneScreen Plus; Du Pont, Boston, MA, USA or BioTrace HP; Gelmen Sciences, Ann Arbor, MI, USA
Multiprime DNA labelling system (Code RPN.1600Z); Amersham
salmon sperm DNA; Sigma
sodium dodecyl sulphate (SDS):- ultrapure for DNA extractions; BDH Chemicals, Poole, England or low grade for washing solutions; Sigma
tetracycline; Sigma
X-ray film (Fuji-RX and Fuji-NC);Fuji Photo Film Co., Tokyo, Japan

All other standard laboratory reagents were purchased from Sigma, BDH or Ajax Chemicals, Auburn, Australia and were of A grade or analytical quality.

2.2.2 Enzymes

The following restriction endonucleases were used:-

1) *BamH*I; Pharmacia
2) *Cfo*I; Promega, Madison, WI, USA
3) *Dra*I; New England Biolabs, Beverly, MA, USA
4) *Msp*I; Boehringer, Mannheim, West Germany
5) *Taq*I; New England Biolabs, Promega
6) *Pst*I; Pharmacia
7) *Pvu*II; Pharmacia
8) *Sal*I; Boehringer
9) *Xba*I; Pharmacia

Other enzymes used in the preparation of genomic DNA and in the radioactive labelling of plasmid DNA were as follows;-

1) Proteinase K; Sigma (12.2 U/mg) or Applied Biosystems (130U/ml)
2) *Escherichia coli* DNA polymerase I (0.5U/μl); Amersham
3) *E. coli* 'Klenow' fragment of DNA polymerase I (1.0U/μl); Amersham
4) *Thermus aquaticus* (*Taq*) DNA polymerase; Perkin Elmer Cetus (5U/μl), Norwalk, CT, USA or Promega (4.5U/μl)

2.2.3 Bacterial growth media

E. coli strains were grown in either LB (Luria-Bertani) medium (per litre): bacto-tryptone 10g; bacto-yeast extract 5g; NaCl 10g; pH 7.5 (Maniatis *et al.*, 1982) or TB (Terrific Broth) medium (per litre): 900ml containing 4ml glycerol; 12g bacto-tryptone; 24g bacto-yeast extract; 100ml of 0.17M KH_2PO_4 and 0.72M K_2HPO_4 (Tartof and Hobbs, 1987).

2.2.4 Bacterial strains and plasmids

The majority of plasmids used in the preparation of cDNA probes were supplied by the American Type Culture Collection. These included pAI-113 (Apolipoprotein A1; APOA1) in pKT218 (Breslow *et al.*, 1982*b*); pAII-E9 (Apolipoprotein A2; APOA2) in pKT218 (Knott *et al.*, 1984*a*; Knott *et al.*, 1984*b*); pUCI-A4 (Apolipoprotein C1; APOC1) (Knott *et al.*, 1984*a*; Knott *et al.*, 1984*c*); pCII-711 (Apolipoprotein C2; APOC2) in pKT218 (Jackson *et al.*, 1984) and pLDLR3 (Low Density Lipoprotein Receptor; LDLR) in pcDV1 (Yamamoto *et al.*, 1984; Hobbs *et al.*, 1985; Hobbs *et al.*, 1987).

pAPOD.6 (Apolipoprotein D; APOD) in pUC18 was the gift of Dr D. Drayna, Genentech, CA, USA (Drayna *et al.*, 1986; Drayna *et al.*, 1987*b*). LPL35 (Lipoprotein Lipase; LPL) was the gift of Dr M. C. Schotz, Lipid Research, UCLA, CA, USA (Wion *et al.*, 1987).

pB8 and pB24 (Apolipoprotein B; APOB; ATCC; Huang *et al.*, 1985) and Probes IV and V (Apolipoprotein; APOA4; a gift from Dr. S.K. Karathanasis; Karathanasis, 1985) were used in preliminary testing on Nauruan samples. A decision was made to limit the number of lipoprotein genes investigated to eight, and these probes were not used in further experiments.

2.3 BLOOD SAMPLES

2.3.1 Micronesian samples

Peripheral blood samples from patients with NIDDM, individuals with impaired glucose tolerance and individuals with normal glucose tolerance were collected during the 1982 and 1987 Nauru Island Surveys (Zimmet *et al.*, 1984; Collins *et al.*, 1989; refer Section 2.1 for Survey details). The blood samples collected in the 1982 survey were centrifuged and separated into plasma, buffy coat and packed red cells. These samples were frozen and transported to the Lions-International Diabetes Institute and/or the John Curtin School of Medical Research, where they were stored in liquid nitrogen until required. Blood samples from the 1987 survey were centrifuged and the plasma separated from the red cells and buffy coat. Approximately

4-6ml of the mixed red blood cells and buffy coats were transported frozen to the John Curtin School of Medical Research and stored at -20°C until required.

2.3.2 Caucasoid control samples

Peripheral blood was collected from healthy laboratory staff, centrifuged and the buffy coats removed. Buffy coat samples were either used immediately for DNA extraction or stored at -20°C until required.

2.4 LABORATORY METHODS

2.4.1 Routine laboratory procedures

All solutions and plastic disposable labware requiring sterilization were autoclaved at 121°C/100kPa for 15min. Glassware was baked at 180°C for at least four hours. Heat labile compounds, such as antibiotics (ampicillin, tetraclycline and chloamphenicol) 2-mercapto-ethanol and 2-dimethyl sulphate were added to solutions after autoclaving. Distilled and deionized water (ddH_2O) was used for all solutions.

Nucleic acids were purified by successive extractions with equal volumes of phenol (saturated with TE buffer: 10mM Tris-HCl; 1mM EDTA; pH 8.0) followed by equal volumes of phenol/chloroform/isoamyl alcohol (25:24:1) followed by equal volumes of chloroform/isoamyl alcohol (24:1) (modified after Maniatis *et al.*, 1982). Ethanol precipitations were carried out with the addition of 0.1 volumes of 5M NaCl or 3M NaOAc and 2 volumes of absolute

alcohol and incubation at -20°C overnight or with the addition of 0.5 volumes of 7.5M NH_4OAc and 2.5 volumes of absolute ethanol and immediate precipitation (Crouse and Amorese, 1987). DNA concentrations and purities were quantified by measuring the optical density of solutions at 260nm and 280nm (1 OD_{260} unit = 50μg/ml for double stranded DNA and 33μg/ml for single stranded DNA; estimates of purity given by $1.8 \leq OD_{260}/OD_{280} \leq 2.0$) (Maniatis *et al.*, 1982).

2.4.2 Biological containment and radiation safety

All recombinant DNA techniques were performed under C1 biological containment conditions and according to the Australian Recombinant DNA Monitoring Committee. Radioactive procedures were performed, and wastes disposed of, under standards set by the Australian National University Radiation Safety Handbook.

2.4.3 Genomic DNA extraction

Genomic DNA was extracted manually from frozen pure buffy coats (1982 samples) and mixed red blood cells and buffy coats (1987 samples) following the method of Grunebaum *et al.* (1984). The 1982 buffy coat samples were thawed for four hours on ice (or overnight at 4°C), washed three times with $T_{20}E_5$ (20mM Tris-HCl; 5mM EDTA; pH 7.5) and incubated at 37°C overnight in 4ml $T_{20}E_5$ in the presence of 200μl 10% ultrapure SDS and 400μl Proteinase K (1mg/ml). Phenol extraction and ethanol precipitation methods have been described elsewhere (Section 2.4.1).

Mixed red blood cell and buffy coat samples collected in the 1987 survey were persistently difficult to extract due primarily to the high proportion of red blood cells present. As many as eight initial $T_{20}E_5$ washes were required and the amount of Proteinase K in the overnight incubation was at times increased twofold. The most intractable samples required up to three phenol, three phenol/chloroform/isoamyl alcohol and three chloroform/isoamyl alcohol extractions. As a consequence yields were sometimes poor.

The Applied Biosytems 340A Nucleic Acid Extractor was used to extract DNA from approximately half of the 1987 survey samples. The Standard Buffy Coat Protocol for 30ml vessels was used with the following modifications implemented to overcome problems caused by the high proportion of red cells:-

1) Half the recommended original sample volume was used
2) Samples were washed five times (instead of twice) in $T_{20}E_5$ before application to the extractor
3) An additional phenol extraction stage was added to the protocol

Buffy coat samples from Caucasoid controls were extracted using the standard procedures described for the 1982 Micronesian samples.

All DNA extracted on the automated extractor was resuspended from precipitate papers in an original volume of 300μl TE with overnight incubation at 55°C. All samples extracted by hand were resuspended in volumes of 50-200μl TE with overnight incubation at 4°C. All samples were

finally suspended at concentrations of 0.3-2.0μg/μl and stored in sealed eppendorf tubes at 4°C.

2.4.4 Polymerase chain reaction amplification

Genomic DNA was amplified by polymerase chain reaction (Saiki *et al.*, 1985) in an Innovonics Gene Machine (Victoria, Australia) thermocycler. A 234bp fragment of APOE exon 4 was amplified using two primers of 22 and 20 bp respectively. The sequence for primer one (designated APOE3 5′ -GGC ACG GCT GTC CAA GGA GC- 3′) was taken from primer PCRE2 of Weisgraber *et al.* (1988) and for primer two (designated SS10 5′ -TCG CGG GCC CCG GCC TGG TAC A- 3′) was modified after primer F4 of Emi *et al.* (1988).

The protocol set out by Hixson and Vernier (1990) was followed with some modifications. A 100μl reaction contained 1.0 μg of genomic DNA, 7.5ng/μl of each primer, 10% dimethyl sulphate (DMSO), 2.5U Taq DNA polymerase, 0.2M of each dNTP and either Cetus Recommended Buffer (CRB:- 50mM KCl; 10mM Tris-HCl, pH 8.3; 1.5mM $MgCl_2$; 0.01% (w/v) gelatin) or Promega supplied *Taq* DNA polymerase buffer (as for CRB with 0.1% Triton X-100 and Tris-HCl at pH 9.0). The inclusion of DMSO in the reaction mix was crucial for amplification of the GC-rich APOE exon 4. Each reaction mix was overlaid with 75μl mineral oil to prevent evaporation and subjected to an initial denaturation of 5min at 95°C, followed by 30 cycles of amplification by primer annealing at 60°C for 1min, extension at 70°C for 1min and denaturation at 95°C for 2min. The protocol was further modified in the course of the study by decreasing the

amount of *Taq* to 1U, raising the extension temperature to 72°C and shortening the extension time to 1min.

Following amplification 1/10th of the PCR product was removed from beneath the mineral oil and visualized on 2% agarose (refer Section 2.4.6). The remainder of the product was frozen at -20°C until required for amplified fragment length polymorphism (AFLP) typing. The presence of contamination was monitored by the inclusion of a negative control in every set of amplifications. This control contained all reaction components excluding genomic DNA.

2.4.5 Restriction endonuclease digestion

For restriction fragment length polymorphism (RFLP) analyses genomic samples containing 8-12μl of DNA were digested with between 40 and 50U of the appropriate enzyme (*Dra*I, *Msp*I, *Taq*I or *Pvu*II) according to the manufacturers recommendations. One fifth of each PCR product was digested with 5U of *Cfo*I for AFLP typing. Stock 10X buffers were used as supplied by the manufacturer or prepared according to the suppliers specifications. All digestions were incubated overnight at 37°C, with the exception of *Taq*I digestions which were incubated for 2hrs at 65°C. Digestions were terminated at completion by immediate ethanol precipitation at room temperature. In the event of overnight (-20°C) precipitation EDTA was added at 25mM/μl reaction volume. All samples were air dried or desiccated and resuspended in 15μl TE.

2.4.6 Agarose gel electrophoresis

Digested genomic DNA was fractionated in 200x200x7mm 0.8-1.0% (w/v) agarose gels in TAE or TBE buffers (TAE: 0.04M Tris-Acetate; 0.001M EDTA; pH 8.0. TBE: 0.089M Tris-borate; 0.089M boric acid; 0.002M EDTA pH 8.0 (Maniatis *et al.*, 1982). Prior to electrophoresis 15% Ficoll loading buffer containing bromophenol blue (0.25%) and xylene cyanol (0.25%) (Maniatis *et al.*, 1982) was added at 3µl per 15µl sample. Immediately before loading the samples were heated at 65°C for 10min and quick-chilled to dissociate annealed compatible ends. These large agarose gels were run using a submarine apparatus (GNA-200, Pharmacia) at 90V for 30min, to run the DNA samples out of the wells, and then set to run at 25V for at least a further 16hrs.

Small mini-gels (100x65x5mm) were used to visualize both PCR product DNA and plasmid DNA. Submarine agarose gels (from 1.0-2.0% (w/v)) were run on a Miniphor Submarine Electrophoresis Unit (LKB 2013, Hoefer, San Francisco, CA, USA) at 100V for 1-2hrs. Samples were treated prior to loading as outlined in the previous paragraph.

Gels were stained in a 1µg/ml ethidium bromide/buffer solution for 15-30min, visualised and photographed under UV light (Sharp *et al.*, 1973).

2.4.7 Southern transfer

Fragmented genomic DNA was transferred from the agarose gels to nylon membranes (GeneScreen Plus or BioTrace HP) following the method of Southern (1975) as modified by Reed

and Mann (1985). Gels were depurinated prior to transfer in 0.25M HCl for 8-15min to aid in the transfer of large fragments. BioTrace HP membranes were saturated in methanol prior to application to the gel; GeneScreen Plus membranes were saturated in ddH_2O. All gels were blotted overnight and restained with ethidium bromide after blotting to check for adequate transfer of fragments. Membranes were rinsed after blotting in 0.2M Tris-HCl pH 7.5, 2XSSC (1XSSC = 150mM NaCl; 15mM trisodium citrate) for 30sec and air dried. The re-hybridization potential of the membrane was improved by baking for 2hr at 80°C and exposure to UV (254nm) for 5min. Nylon membranes were stored in Glad Wrap, in the dark at room temperature.

2.4.8 Polyacrylamide gel electrophoresis

All *Cfo*I digested APOE PCR products pretreated as above (Section 2.4.6) were fractionated on 160x160x1.5mm 10% polyacrylamide gels using Hoefer supplied plates and tank apparatus (Vertical Slab Gel Electrophoresis Unit, SE 600, Hoefer). The 10% gels were prepared as follows:-

10ml	40% polyacrylamide
5ml	10xTBE stock buffer
34ml	ddH_2O
500µl	10% ammonium persulphate
50µl	99% TEMED

Stock solutions of 40% polyacrylimide (acrylamide/N,N'-methylene-*bis*-acrylamide (38/2, w/w) were prepared, deionized, filtered, de-gased and stored at 4°C in the dark. Ammonium persulphate (10%) was made weekly and stored at 4°C. Gels were run in TBE, at 30mA/gel for approximately

3hr before staining in ethidium bromide (1μg/ml buffer) for 30min and visualization under UV (Maniatis *et al.*, 1982).

2.4.9 Transformation of competent cells

Competent cells were prepared using the method of Dagert and Ehrlich (1977), snap frozen and stored at -70°C. Transformations of *E. coli* strains with plasmid were performed according to Yanisch-Perron *et al.* (1985) and grown in LB.

2.4.10 Preparation of plasmid DNA

Plasmid DNA was prepared according to the alkaline lysis method (Ish-Horowicz and Burke, 1981; Maniatis *et al.*, 1982). Large scale preparations were made in 500ml cultures of LB or TB with ampicillin (50μg/ml) or tetraclycline (15μg/ml) and amplification overnight with chloramphenicol (136μg/ml) after reaching an OD_{600} of 0.7. Small preparations were made from 1ml of 5ml overnight LB cultures without amplification. Plasmid DNA was resuspended after CsCl purification (Maniatis *et al.*, 1982) in TE to a final stock concentration of ≥1μg/μl and stored at -70°C in 50μl aliquots. Working solutions of 0.25μg/μl were stored at 4°C.

2.4.11 Recovery of DNA from agarose gels

A number of the cDNA probes used labelled and hybridized more efficiently when the insert was removed from the host plasmid. Plasmid DNA was digested and the fragments run on agarose gels. A maximum agarose fragment of 7x5x5mm

containing the appropriate insert was cut from the gel after staining and the DNA electo-eluted at 100V for 40min using a IBI Model UEA Electroeluter (Bromma, Sweden). Inserts were ethanol precipitated and resuspended in TE.

2.4.12 Radioactive labelling of plasmid DNA

cDNA probes were prepared from whole plasmids or inserts by nick translation (Kelly *et al.*, 1970) or random priming (Feinberg and Vogelstein, 1983, 1984) respectively. Nick translation and random priming reactions were performed using the Amersham Nick Translation Kit and the Amersham Multiprime DNA Labelling Systems according to the suppliers specifications. The reaction mixes were as follows:-

<u>Nick translation</u>
0.25μg plasmid DNA
10.0μl mixed dNTP buffers
(dCTP, dGTP, dTTP each at 300μM)
2.5μl [α-^{32}P]dATP
5.0μl DNA polymerase I (2.5U)
xμl ddH_2O to 50μl

<u>Random priming</u>
0.25μg insert DNA
(pre-boiled for 2min and kept on ice)
4.0μl each of dCTP, dGTP, dTTP as supplied
2.5μl [α-^{32}P]dATP
5.0μl 10x supplied reaction buffer
5.0μl 10x supplied hexanucleotide primer buffer
2.0μl DNA polymerase I 'Klenow' fragment (2U)
xμl ddH_2O to 50μl

The reactions were incubated at 15°C for 1hr and 37°C for 30min respectively. The labelled probes were ethanol precipitated, resuspended in at least 200μl TE and stored on ice until required.

2.4.13 Hybridizations

The protocol of Nasmyth (1982) was followed to hybridize labelled cDNA probes to the nylon membranes. Two membranes were placed, DNA side out, in plastic bags and prehybridized at room temperature for at least 2hr in a prehybridization mix of 0.46 volumes ddH_2O containing 50μg/ml preboiled sonicated salmon sperm DNA and Nasmyth Solution to a total volume of 20ml. The Nasmyth solution (1.1M NaCl; 0.333M NA_2HPO_4; 0.111M EDTA, pH 6.2; 1.85% N-lauroylsarcosine and 18.5% dextran sulphate) was made weekly in small quantities and stored at 4°C.

Hybridizations solutions were identical to prehybridization solutions but also included 0.05μg ($1x10^8$ cpm/μg) of labelled probe per 100cm^2 of membrane (and per 5ml hybridization solution). The sonicated salmon sperm DNA and the labelled probe were boiled for 10min and chilled on ice before addition to the hybridization bag. Membranes were incubated at 65°C for 16-48hr in a shaking water bath.

Membranes were put through a series of washes after hybridization with the aim of removing non-specific hybridization and lowering background radioactivity. An initial room temperature wash of 5min in 2xSSC, 0.1% SDS was followed by consecutive washes of 30min in 0.1% SDS and an increasing stringency of 1xSSC, 0.5xSSC and 0.1xSSC. Membranes were checked after each wash, air dried for 5-10min and wrapped in Glad Wrap when they had reached 3-15cps. Fuji Xray film was exposed to the membranes for 3-

10 days at -70°C, in the presence of Lighting Plus intensifying screens (Du Pont) and developed automatically.

Membranes were stripped in 0.4M NaOH at 42°C for 30min and 0.1xSSC, 0.1% SDS, 0.2M Tris-HCl pH 7.5 at 42°C for 30min. Both GeneScreen Plus and BioTrace HP membranes were reused several times and were stored moist between use.

2.5 STATISTICS AND COMPUTING METHODS

Data management and statistical analyses were carried out on a VAX 8700 supported by the Australian National University Computer Service Centre, and on a NEC PowerMate 386/20 (NEC Inc, Tokyo, Japan). The following software packages were used:- SPSS^{x} Statistical Package for the Social Sciences (SPSS Inc., 1986) for general data manipulation, descriptive statistics, comparison of means (both *t*-test and ANOVA), multiple linear regressions; GLIM Generalized Linear Interactive Modelling (Payne, 1986; Healy, 1988); LIPED for linkage analysis (Ott, 1974); INSTAT (GraphPAD Software, San Diego, CA, USA) for chi-square tests and Fisher's exact probabilities; PC/GENE (IntelliGenetics Inc., Mountain View, CA, USA) for sequence analysis; BIOSYS (Swofford and Selander, 1981) for gene frequency calculations and tests of Hardy-Weinberg equilibrium.

SECTION 3

LIPOPROTEIN GENE VARIATION IN MICRONESIANS: COMPARISONS WITH OTHER ETHNIC GROUPS

3.1 AIMS

1) To establish the frequencies of seven lipoprotein gene RFLPs and of the common apolipoprotein E functional alleles in a Micronesian population.
2) To compare these allele frequencies with known frequencies in other ethnic groups.
3) To estimate the information content of these polymorphisms and their suitability as candidate markers for NIDDM and lipoproteinemia.

3.2 INTRODUCTION

The characterization, localization and sequencing of the genes involved in lipoprotein transport and metabolism has led to many studies of the potential role of these genes in lipoproteinemias and atherosclerosis. As a corollary there is now information available on the frequencies of protein variants and RFLPs for lipoprotein genes in different ethnic groups. However, these data are limited in some cases to Caucasoid populations, and are more extensive for the lipoproteins genes which have been more heavily implicated in the aetiology of lipoproteinemia and atherosclerosis. In particular the APOA1, APOC2 and APOE genes have been examined in a range of ethnic groups whilst there is a paucity of ethnic information for APOA2, APOC1, APOD, LDLR and LPL.

This chapter reports specific RFLP and AFLP frequencies for sites within these eight lipoprotein genes in a Micronesian

population and compares them with allele frequencies in other ethnic groups. In this chapter, and the remainder of the thesis, the APOE alleles will be referred to as APOE*2, APOE*3 and APOE*4, according to convention, while a code of gene name-enzyme initial will be used to distinguish between the different RFLPs. The number 1 will always be used within the enzyme abbreviation of this code to refer to the presence of the polymorphic site and consequently identifies the smallest polymorphic fragment sizes associated with each RFLP. For example APOA1-T1 refers to the presence of the *Taq*I restriction site in APOA1.

3.2.1 Lipoprotein genes in different ethnic groups

3.2.1.1 Apolipoprotein A1

The *Msp*I site within the third intron of APOA1 was first reported by Seilhamer *et al*. (1984). This *Msp*I site has been examined extensively (Table 3.1), as it falls within the APOA1-C3-A4 lipoprotein gene cluster of chromosome 11q23-qter, and has been used, along with a range of other sites, most commonly *Pst*I and *Sst*I, to establish haplotype frequencies in lipoproteinemia and heart disease studies (Anderson *et al*., 1986; Humphries *et al*., 1987; Antonarakis *et al*., 1988; Paulweber *et al*., 1988; Cole *et al*., 1989). There is sufficient variability in allele frequencies within ethnic groups at this locus to make pooling of results from the different reports inappropriate.

Table 3.1 APOA1 *Msp*I allele frequencies in non-Micronesians.

Ethnicity†	APOA1-M1	N	Reference
African	0.74	27	Paul *et al.*, 1987
African American	0.87	62	Thompson *et al.*, 1988
African American	0.96	67	Antonarakis *et al.*, 1988
Asian Indian	0.54	23	Paul *et al.*, 1987
Caucasoid	0.88	70	Coleman *et al.*, 1986
Caucasoid	0.89	231	Thompson *et al.*, 1988
Caucasoid	0.90	92	Ferns and Galton, 1986*a*
Caucasoid	0.93	142	Hegele *et al.*, 1989
Caucasoid	0.94	100	Frossard *et al.*, 1986
Caucasoid	0.95	129	Antonarakis *et al.*, 1988
Caucasoid	0.97	52	Paul *et al.*, 1987
Japanese	0.44	27	Paul *et al.*, 1987
Japanese	0.44	68	Thompson *et al.*, 1988
Japanese	0.60	82	Hamaguchi *et al.*, 1987

† Heterogeneity tests:- African χ^2=19.62, d.f.=2, p<0.001; Caucasoid χ^2=16.23, d.f.=6, p=0.012; Japanese χ^2=8.52, d.f.=2, p=0.014.

3.2.1.2 Apolipoprotein A2

Another *Msp*I site is found within an *Alu* sequence 3′ to the APOA2 gene on chromosome 1q21-q23 (Scott *et al.*, 1985). Reports of frequencies for this site are limited to Caucasoid populations (Table 3.2), among which significant differences in allele frequencies are found.

Table 3.2 APOA2 *Msp*I allele frequencies in non-Micronesians.

Ethnicity†	APOA2-M1	N	Reference
Caucasoid	0.81	87	Scott *et al.*, 1985
Caucasoid	0.81	59	Ferns *et al.*, 1986*b*
Caucasoid	0.85	163	Deeb *et al.*, 1986
Caucasoid	0.85	85	Rajput-Williams *et al.*, 1989
Caucasoid	0.90	160	Hegele *et al.*, 1989

† Heterogeneity test:- Caucasoid χ^2=9.72, d.f.=4, p=0.045.

3.2.1.3 Apolipoprotein C1

Apolipoprotein C1 is situated within the apolipoprotein gene cluster at 19q12-q13.2. A *Dra*I polymorphic site has been reported at this locus in one Caucasoid sample with a frequency of 0.786 (APOC1-D1, N=378, Frossard *et al.*, 1987).

3.2.1.4 Apolipoprotein C2

A *Taq*I polymorphic site is present some 2kb 3′ of the apolipoprotein C2 gene in the chromosome 19 cluster (Humphries *et al.*, 1983). The site has been examined in a number of different ethnic groups, as presented in Table 3.3. APOC2-T1 frequencies for the four Caucasoid populations do not differ significantly from each other and have been pooled.

Table 3.3 APOC2 *Taq*I allele frequencies in non-Micronesians.

Ethnicity†	APOC2-T1	N	Reference
African	0.48	46	Williams *et al.*, 1985
Asian Indian	0.50	50	Williams *et al.*, 1985
Caucasoid	0.35	103	Deeb *et al.*, 1986
Caucasoid	0.40	114	Williams *et al.*, 1985
Caucasoid	0.44	100	Frossard *et al.*, 1987
Caucasoid	0.44	90	Humphries *et al.*, 1983
Pooled Caucasoid	0.41	407	
Chinese	0.58	38	Williams *et al.*, 1985
Japanese	0.56	40	Williams *et al.*, 1985
West Indian	0.52	52	Williams *et al.*, 1985

† Heterogeneity test:-Caucasoid $\chi^2=4.54$, d.f.=3, p=0.210.

3.2.1.5 Apolipoprotein D

A *Taq*I polymorphic site has been reported for apolipoprotein D (Drayna *et al.*, 1987*b*). The frequency of

APOD-T1 in Caucasoids was 0.18 (N=103, Drayna *et al.*, 1987*b*).

3.2.1.6 Apolipoprotein E

Apolipoprotein E gene and protein polymorphisms have been extensively examined. Three common alleles for APOE, APOE*2, APOE*3 and APOE*4, are co-dominantly inherited and code for three apolipoprotein E isoforms. These isoforms differ at amino acid residues 112 and 158. Isoform APOE*2 has cysteine residues at both 112 and 158, APOE*3 has a cysteine at 112 and an arginine at 158, whilst APOE*4 has arginine residues at both sites (Rall *et al.*, 1982). In addition to producing charge differences, these substitutions alter two recognition sites for *Cfo*I and its isoschizomer *Hha*I (Hixson and Vernier, 1990).

The relationships between APOE isoforms and hyperlipoproteinemias or atherosclerosis are now well established (Davignon *et al.*, 1988). The apolipoprotein E isoforms APOE*2, APOE*3 and APOE*4 can be detected by two-dimensional gel electrophoresis (Zannis and Breslow, 1981) or isoelectric focussing (Utermann *et al.*, 1982) with immunoblotting (Kamboh *et al.*, 1988*b*). The corresponding DNA polymorphisms can be detected using a range of techniques utilizing the polymerase chain reaction and either allele specific oligonucleotide probing (Emi *et al.*, 1988; Smeets *et al.*, 1988; Weisgraber *et al.*, 1988; Kontula *et al.*, 1990), oligonucleotide "melting" (Funke *et al.*, 1986), post-amplification cleavage with *Hha*I (or its isoschizomer *Cfo*I) (Hixson and Vernier, 1990; Kontula *et*

al., 1990), allele specific priming (Wenham *et al.*, 1991) or primer guided incorporation assays (Syvanen *et al.*, 1990).

Allele frequencies are known for a number of different ethnic groups (Table 3.4). Although APOE allele frequencies do not differ significantly between the two African samples there is within ethnic group variability in both Caucasoids and Japanese.

Table 3.4 APOE allele frequencies in non-Micronesians.

E*2	E*3	E*4	N	Reference
African†				
0.027	0.677	0.296	365	Sepehrnia *et al.*, 1988a
0.034	0.706	0.260	194	Kamboh *et al.*, 1989
Pooled African				
0.029	0.687	0.284	559	
Amerindian				
0.000	0.816	0.184	95	Asakawa *et al.*, 1985
Caucasoid†				
0.039	0.767	0.194	1157	Lehtimaki *et al.*, 1990
0.041	0.733	0.227	615	Ehnholm *et al.*, 1986
0.076	0.788	0.136	1209	Ordovas *et al.*, 1987
0.077	0.773	0.150	1031	Utermann *et al.*, 1982
0.078	0.770	0.152	102	Sing and Davignon, 1985
0.079	0.784	0.136	624	Lenzen *et al.*, 1986
0.080	0.770	0.150	400	Cumming and Robertson, 1984
0.082	0.751	0.168	2000	Klasen *et al.*, 1987
0.082	0.782	0.136	1557	Assmann *et al.*, 1984
0.12	0.72	0.16	426	Wardell *et al.*, 1982
0.120	0.764	0.116	303	Gueguen *et al.*, 1989
0.130	0.742	0.128	223	Boerwinkle *et al.*, 1987
Chinese				
0.084	0.852	0.064	196	Davignon *et al.*, 1988
Japanese†				
0.023	0.891	0.086	110	Asakawa *et al.*, 1985
0.035	0.872	0.093	188	Kobori *et al.*, 1988
0.037	0.846	0.117	576	Eto *et al.*, 1986a
0.044	0.863	0.093	91	Imari *et al.*, 1988
0.082	0.850	0.068	319	Utermann, 1987
Mayan				
0.000	0.911	0.089	135	Kamboh *et al.*, 1991
Mexican				
0.039	0.859	0.102	963	Hanis *et al.*, 1991
New Guinean				
Pawaia				
0.138	0.603	0.259	58	Kamboh *et al.*, 1990
Huli				
0.154	0.356	0.490	52	Kamboh *et al.*, 1990

† Heterogeneity testing:- African χ^2=1.77, d.f.=2, p=0.412; Caucasoid high^{est} against low^{est} E*2 χ^2=67.19, d.f.=2, p<0.001; Caucasoid high^{est} against low^{est} E*3 χ^2=60.37, d.f.=2, p<0.001; Caucasoid high^{est} against low^{est} E*4 χ^2=114.81, d.f.=2, p<0.001; Japanese χ^2=34.00, d.f.=8, p<0.001.

3.2.1.7 Low density lipoprotein receptor

The most commonly investigated RFLP at the LDLR locus is the result of a *Pvu*II site in intron 15 (Hobbs *et al.*, 1985). This *Pvu*II site has been studied in the general population (Berg, 1990; Humphries *et al.*, 1991) and in pedigrees segregating for familial hypercholesterolemia. In some families affected with familial hypercholesterolemia the *Pvu*II RFLP, along with others (*Nco*I and *Stu*I) were used as genetic markers for identification of at risk individuals (Kotze *et al.*, 1987; Daga *et al.*, 1988; Kotze *et al.*, 1989; Schuster *et al.*, 1989). Reports of allele frequencies for this site have been restricted to Caucasoid populations (Table 3.5). There is a significant difference between the lowest and highest allele frequencies reported for this site.

Table 3.5 LDLR *Pvu*II allele frequencies in non-Micronesians.

Ethnicity†	LDLR-P1	N	Reference
Caucasoid	0.18	239	Pedersen and Berg, 1990
Caucasoid	0.226	96	Humphries *et al.*, 1985
Caucasoid	0.23	324	Schuster *et al.*, 1990
Caucasoid	0.236	19	Hobbs *et al.*, 1985
Caucasoid	0.29	85	Hegele *et al.*, 1990
Caucasoid	0.32	289	Humphries *et al.*, 1991
Caucasoid	0.32	60	Kotze *et al.*, 1987
Caucasoid	0.402	112	Daga *et al.*, 1988

† Heterogeneity test: Caucasoid high^{est} against low^{est} χ^2=38.80, d.f.=1, p<0.001.

3.2.1.8 Lipoprotein lipase

Lesions in the LPL gene have been found in pedigrees segregating for lipoprotein lipase deficiency (Devlin *et*

al., 1990; Emi *et al.*, 1990*a*, 1990*b*; Hata *et al.*, 1990*b*). However the lipoprotein lipase locus has not been studied extensively with regard to lipoprotein variation in the general population. Allele frequency estimation of the *Pvu*II site in intron 6 (Fisher *et al.*, 1987) is limited to Caucasoid populations (Table 3.6). The frequency of LPL-P1 does not differ significantly among the three Caucasoid populations.

Table 3.6 LPL *Pvu*II allele frequencies in non-Micronesians.

Ethnicity†	LPL-P1	N	Reference
Caucasoid	0.541	186	Chamberlain *et al.*, 1989
Caucasoid	0.554	122	Thorn *et al.*, 1990
Caucasoid	0.59	49	Fisher *et al.*, 1987
Pooled Caucasoid	0.552	357	

† Heterogeneity test: Caucasoid χ^2=0.84, d.f.=2, p=0.659.

3.3 MATERIALS AND METHODS

3.3.1 Sample selection

A series of samples, unrelated by surname, were selected randomly from the 1987 Nauru Survey blood sample collections without reference to diabetic status. Only respondents claiming Nauruan ethnicity were included, although this does not preclude the possibility of ancestral foreign admixture. Caucasoid samples, to be used as experimental controls on hybridization filters, were obtained from healthy laboratory staff.

3.3.2 Sample description

A total sample of 119 was obtained comprising 67 females and 52 males, with a mean age of 40.59 years (SE 1.25).

Mean age did not differ significantly between females and males (Female mean 39.55 years, SE 1.64; Male mean 41.92, SE 1.93; t=0.94, d.f.=117, p=0.348).

3.3.3 APOA1, APOA2, APOC1, APOC2, APOD, LDLR and LPL RFLP typing

Genomic DNA was prepared from the samples, digested with *Dra*I, *Msp*I, *Pvu*II and *Taq*I and transferred to nylon membranes as described in Chapter 2. The cut DNA was hybridized to plasmids containing the cDNA probes discussed in Section 2.2.4 (Table 3.7).

3.3.4 APOE AFLP typing

A 234bp fragment of APOE exon 4 was amplified from each sample and digested with *Cfo*I according to the methods described in Chapter 2.

3.3.5 Statistical methods

Sample allele frequencies were calculated by counting and standard errors for these frequencies calculated using the normal approximation to the binomial distribution, where SE = $\sqrt{pq/2n}$ (p = frequency of allele 1, q = frequency of allele 2, n = number of individuals) (Li, 1955). All 2x2 χ^2 goodness of fit tests were performed with continuity correction. In 2x2 comparisons, where cell sizes were less than five, Fisher's exact test was used (Snedecor and Cochran, 1980). In tests of Hardy-Weinberg equilibrium exact significance probabilites (analagous to Fisher's test) were calculated when cell sizes were less than five (Swofford and Selander, 1981). No adjustments were made for

because such adjustments conservatively reduce the type 1 error for null associations but increase the type II error.
multiple comparisons (Rothman, 1990). The polymorphic information content (PIC) for each polymorphism detected in the Micronesian population, and for the polymorphisms reviewed in Section 3.2, were calculated according to Botstein *et al.* (1980).

Table 3.7 Probe descriptions, enzyme combinations and fragment sizes.

Gene	Probe†	Size (bp)	Vector	Enzyme	Fragment (kb)
APOA1	pAI-113	600	pKT218	*Msp*I	M1 1.0 M2 1.7
APOA2	pAII-E9	430	pKT218	*Msp*I	M1 3.0 M2 3.7
APOC1	pUCI-A4	420	unknown	*Dra*I	D1 2.6/7.6 D2 10.2
APOC2	pCII-711	500	pKT218	*Taq*I	T1 3.5 T2 3.8
APOD	pAPOD.6	740	pUC18	*Taq*I	T1 2.2 T2 2.7
LDLR	pLDLR-2HHI	1900	§	*Pvu*II	P1 14.0 P2 16.5
LPL	pLPL35	2413	unknown	*Pvu*II	P1 2.7/4.9 P2 7.6

† Refer to Sections 2.2.4 and 3.2 for probe source and references.
§ pLDLR-2HHI cut from pLDLR3 (5.3kb probe in pcDV1).

3.4 RESULTS

3.4.1 APOA1, APOA2, APOC1, APOC2, APOD, LDLR and LPL RFLP typing

Hybridization of *Msp*I digested genomic DNA with the 600bp plasmid probe pAI-113 resulted in a 1.7kb fragment in the presence of the *Msp*I site, and a 1.0kb fragment in its absence (Figures 3.1a and 3.2). These alleles will be referred to as APOA1-M1 (1.0kb) and APOA1-M2 (1.7kb). Digestion with *Msp*I and hybridization with the 600bp plasmid pAII-E9 yielded a 3.0kb fragment (APOA2-M1) in the

presence of the APOA2 *Msp*I site and a 3.7kb fragment (APOA2-M2) in its absence (Figures 3.1b and 3.3).

After *Dra*I digestion of genomic DNA, hybridization with the pUC1-A4 probe for APOC1 detected monomorphic fragments of 3.5 and 7.0kb and a single two-allele polymorphism with bands at 2.6 and 7.6kb (APOC1-D1) or 10.2kb (APOC1-D2) (Figure 3.4).

The diallelic *Taq*I site in APOC2 resulted in bands of either 3.5kb (APOC2-T1) or 3.8kb (APOC2-T2) when *Taq*I digested genomic DNA was probed with pUCII-117 (Figures 3.1c and 3.5). Digestion of genomic DNA with *Taq*I and subsequent probing with pAPOD.6 detected a monomorphic fragment of 3.2kb and diallelic bands at 2.2kb (APOD-T1) and 2.7kb (APOD-T2) (Figure 3.6).

When *Pvu*II digested genomic DNA was probed with pLDLR-2HHI (a 1.9kb *Bam*HI fragment of pLDLR3) the presence of the *Pvu*II site was revealed by a 14.0kb fragment (LDLR-P1), and its absence by a 16.5kb fragment (LDLR-P1). pLDLR-2HHI also detected an invariant band of 3.6kb (Figures 3.1d and 3.7). Invariant bands at 0.8, 1.9, 2.9, 3.2 and 6.6kb, and a diallelic polymorphism of 7.6kb (LPL-P2) or 2.7 and 4.9kb (LPL-P1) were detected after hybridization of *Pvu*II digested genomic DNA to pLPL35 (Figures 3.1e and 3.8).

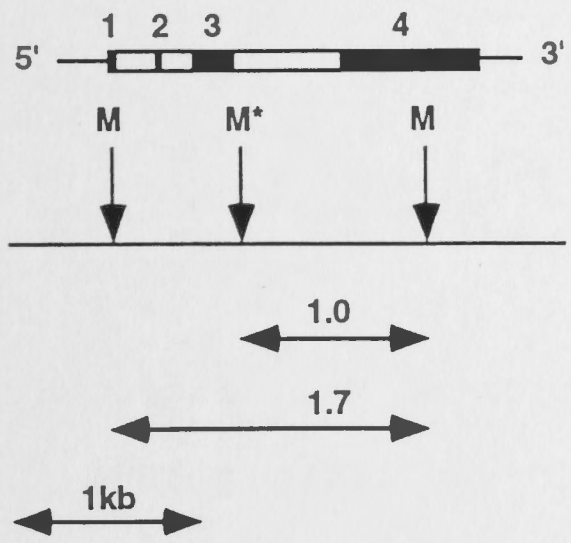
3.4.2 APOE AFLP typing

Six *Cfo*I restriction sites were present in the amplified APOE*4 sequence. One of these sites was absent in APOE*3 and two were absent in APOE*2 (Figure 3.9a). Consequently

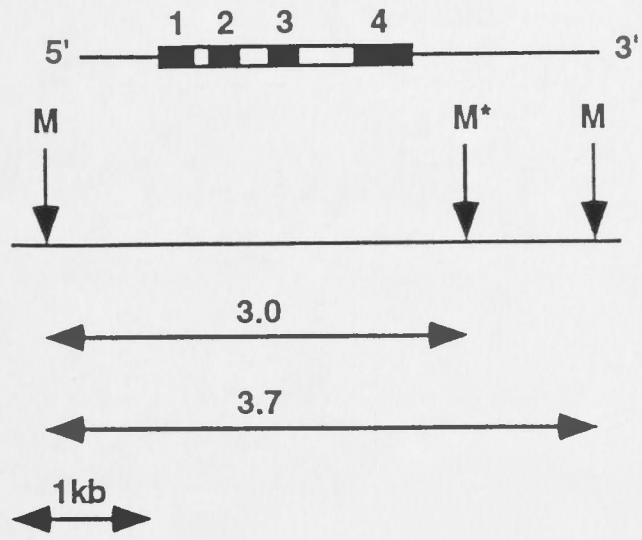
unique combinations of *Cfo*I fragment sizes were present in the digested amplification products of all six genotypes (Figure 3.9a). The fragments revealed upon digestion of the amplification products of homozygotes are presented in Figure 3.9b.

Figure 3.1 Restriction maps of apolipoprotein and lipoprotein restriction fragment length polymorphisms. **a** APOA1; **b** APOA2; **c** APOC2, **d** LDLR; **e** LPL. Solid bars = exons; open bars = introns; * = polymorphic site; M = *Msp*I; T = *Taq*I; P = *Pvu*II.

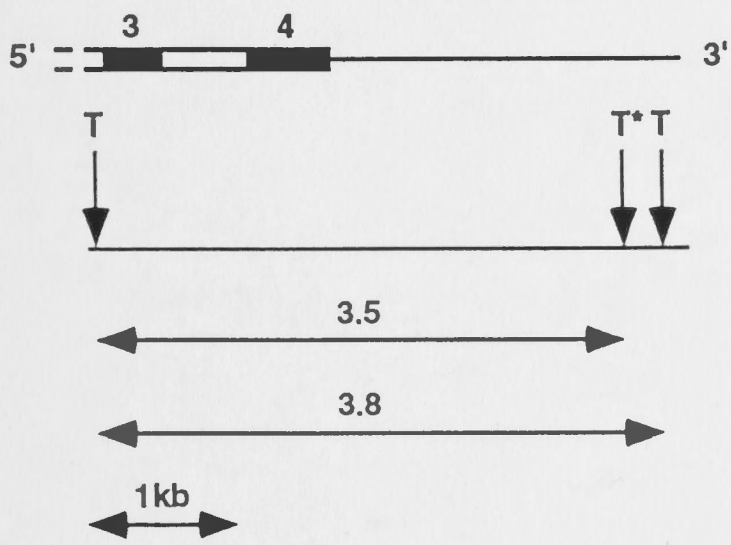
a APOA1



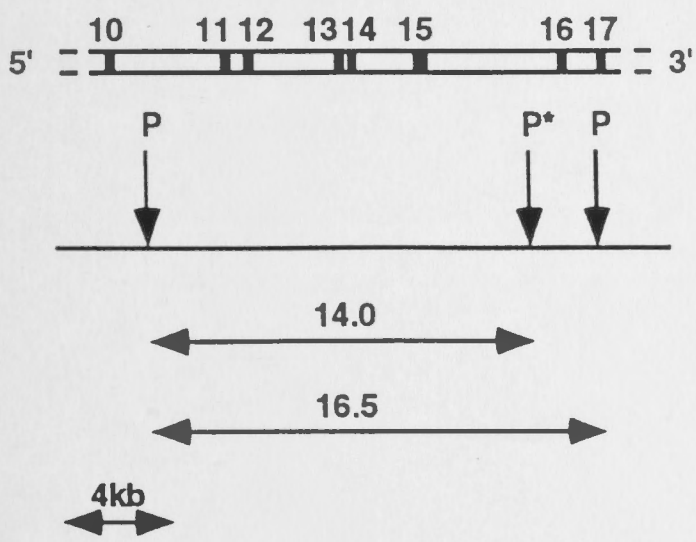
b APOA2



c APOC2



d LDLR



e LPL

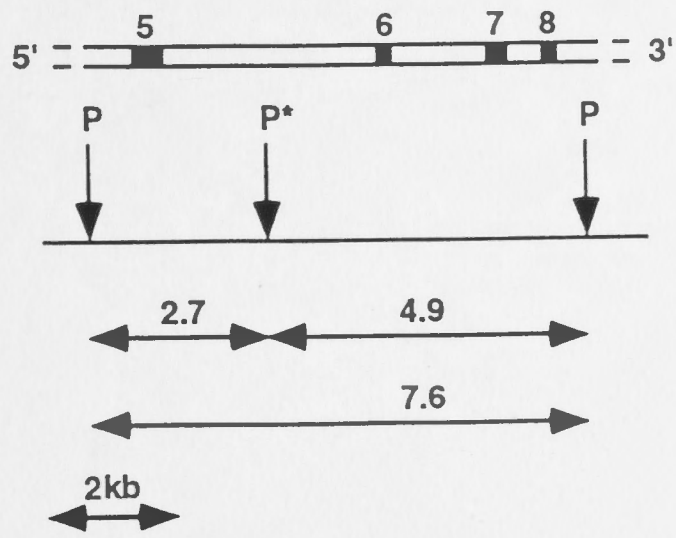


Figure 3.2 APOA1(pAI-113) hybridization of *Msp*I digested genomic DNA. Lane 1 heterozygote (1.0 and 1.7kb); lane 2 APOA1-M2 homozygote (1.7kb); lane 3 APOA1-M1 homozygote (1.0kb).

APOA1(pA1-113)x*Msp* I

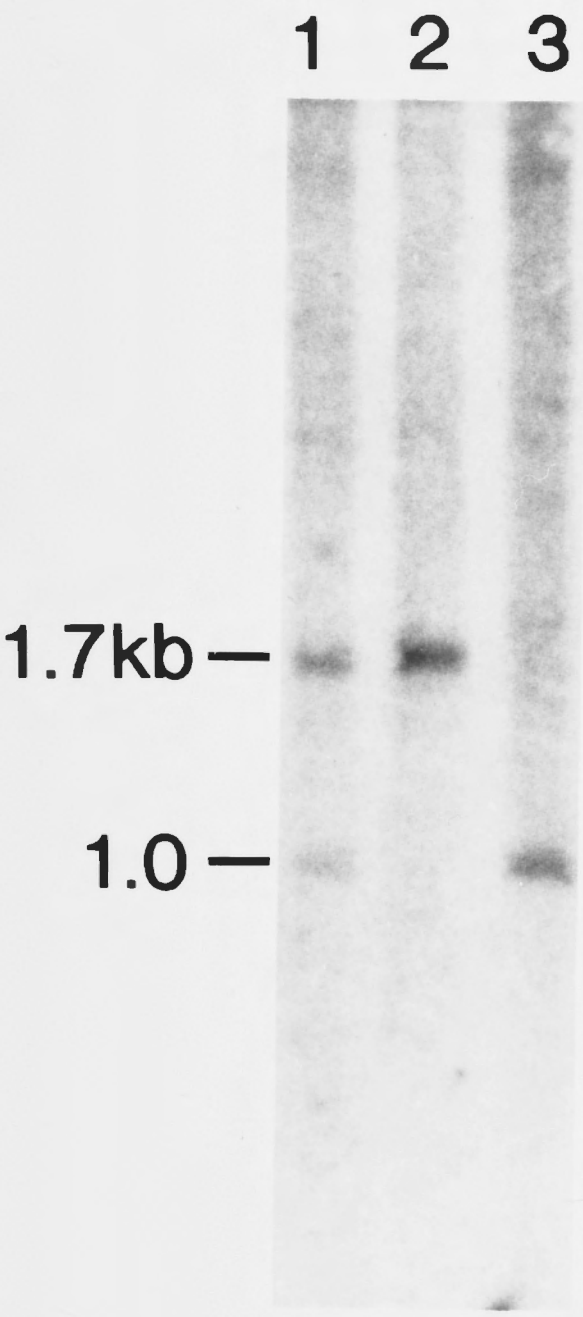


Figure 3.3 APOA2(pAII-E9) hybridization of *Msp*I digested genomic DNA. Lane 1 APOA2-M1 homozygote (3.0kb); lane 2 heterozygote (3.0 and 3.7kb); lane 3 APOA2-M2 homozygote (3.7kb).

APOA2(pAII-E9)x*Msp* I

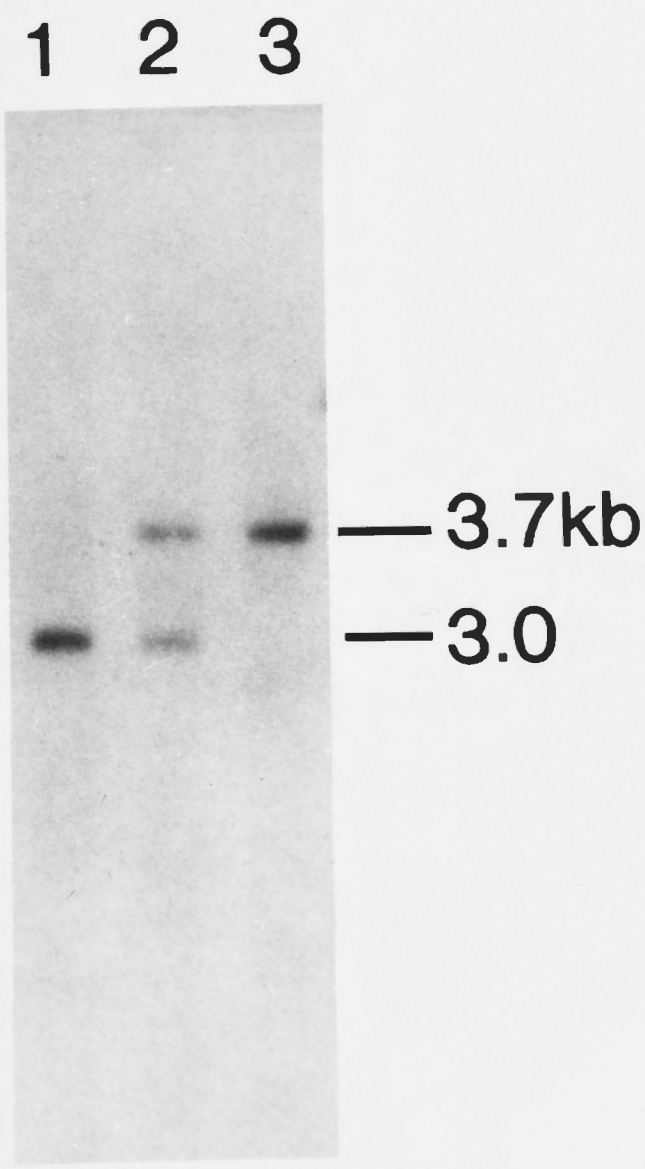


Figure 3.4 APOC1(pUCI-A4) hybridization of *Dra*I digested genomic DNA. Lane 1 heterozygote (2.6/7.6 and 10.2kb); lanes 2 and 3 APOC1-D1 homozygotes (2.6/7.6kb); lane 4 APOC1-D2 homozygote (10.2kb).

APOC1(pUCI-A4)x*Dra* I

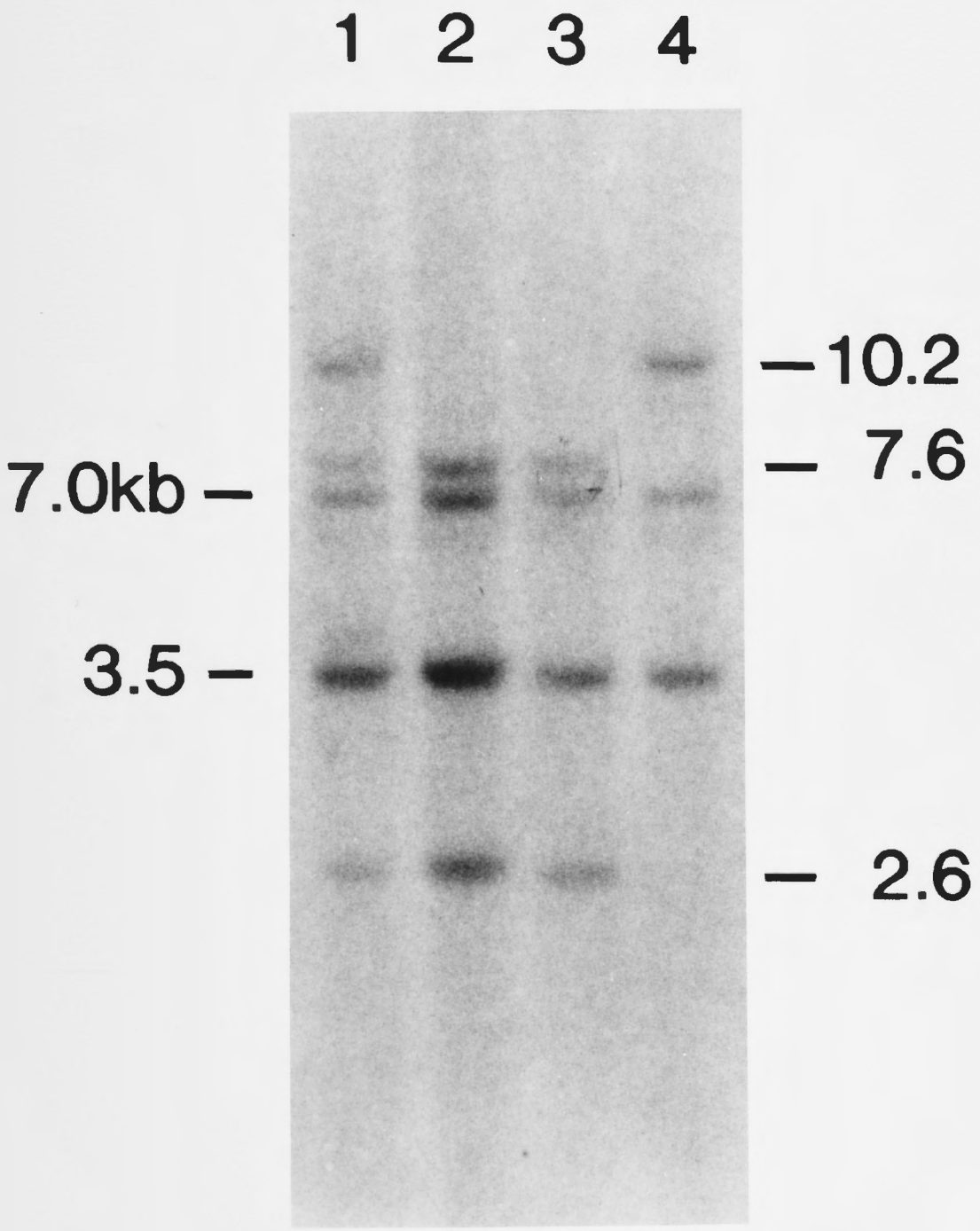


Figure 3.5 APOC2(pCII-711) hybridization of *Taq*I digested genomic DNA. Lane 1 APOC2-T1 homozygote (3.5kb); lane 2 APOC2-T2 homozygote (3.8kb); lane 3 heterozygote (3.5 and 3.8kb).

APOC2(pCII-711)x *Taq* I

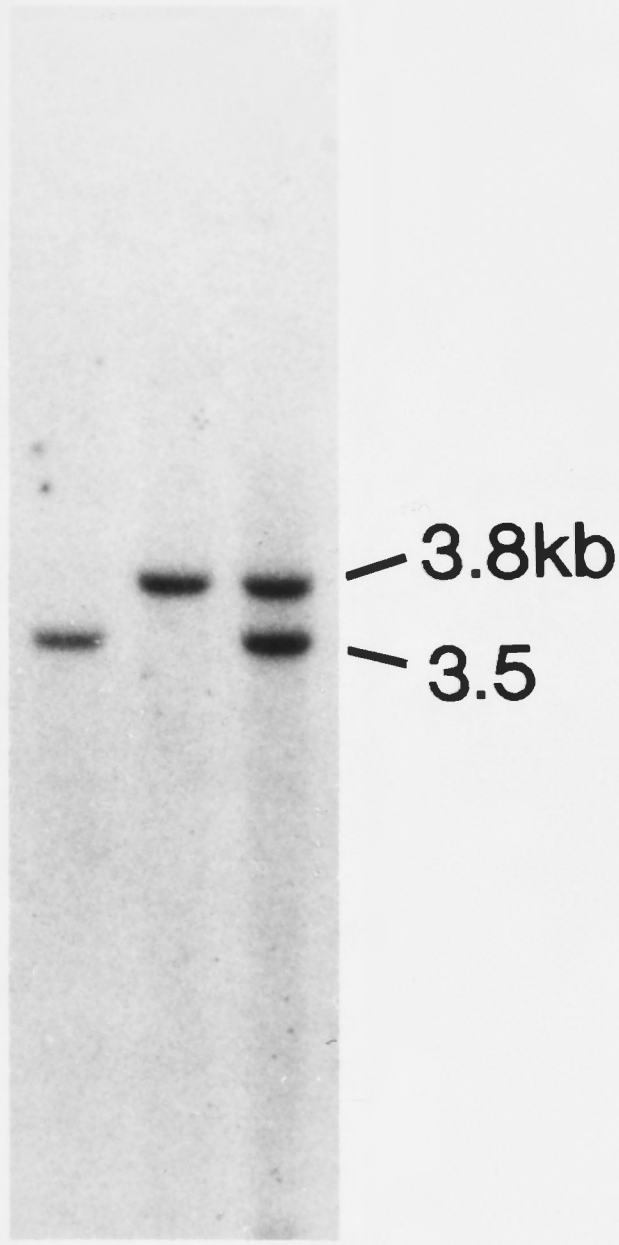


Figure 3.6 APOD(pAPOD.6) hybridization of *Taq*I digested genomic DNA. Lane 1 APOD-T2 homozygote (2.7kb); lanes 2 and 3 heterozygotes (2.2 and 2.7kb); lane 4 APOD-T1 homozygote (2.2kb).

APOD(pAPOD.6)x *Taq* I

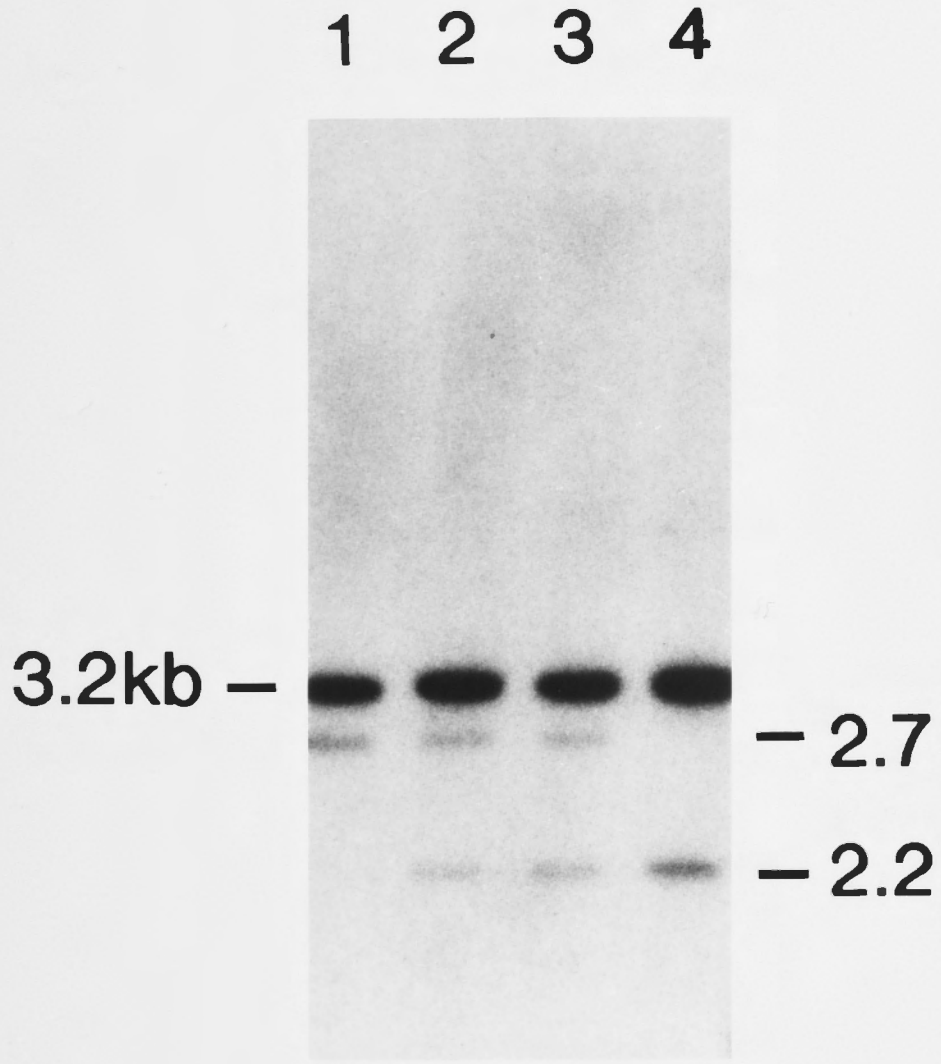


Figure 3.7 LDLR(pLDLR-2HHI) hybridization of *Pvu*II digested genomic DNA. Lane 1 heterozygote (14.0 and 16.5kb); lane 2 LDLR-P1 homozygote (14.0kb); lane 3 LDLR-P2 homozygote (16.5kb).

LDLR(pLDLR-2HHI)x*Pvu* II

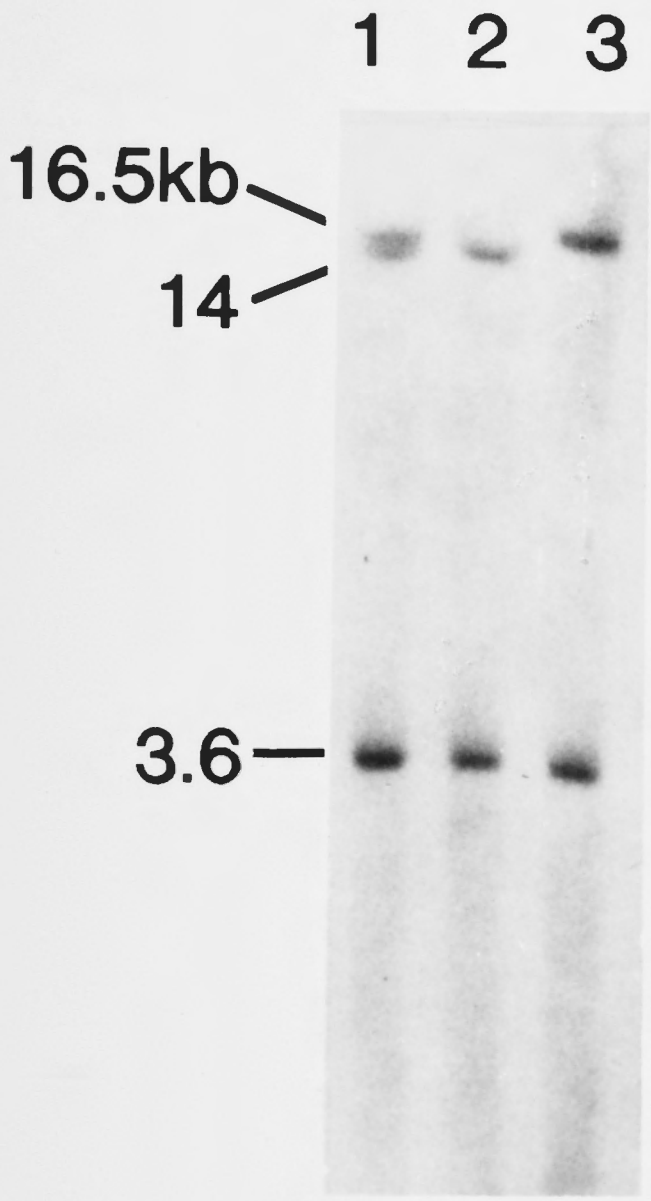


Figure 3.8 LPL(pLPL35) hybridization of *Pvu*II digested genomic DNA. Lane 1 LPL-P1 homozygote (2.7/4.9kb); lane 2 heterozygote (2.7/4.9 and 7.6kb); lane 3 LPL-P2 homozygote (7.6kb).

LPL(pLPL35)x*Pvu* II

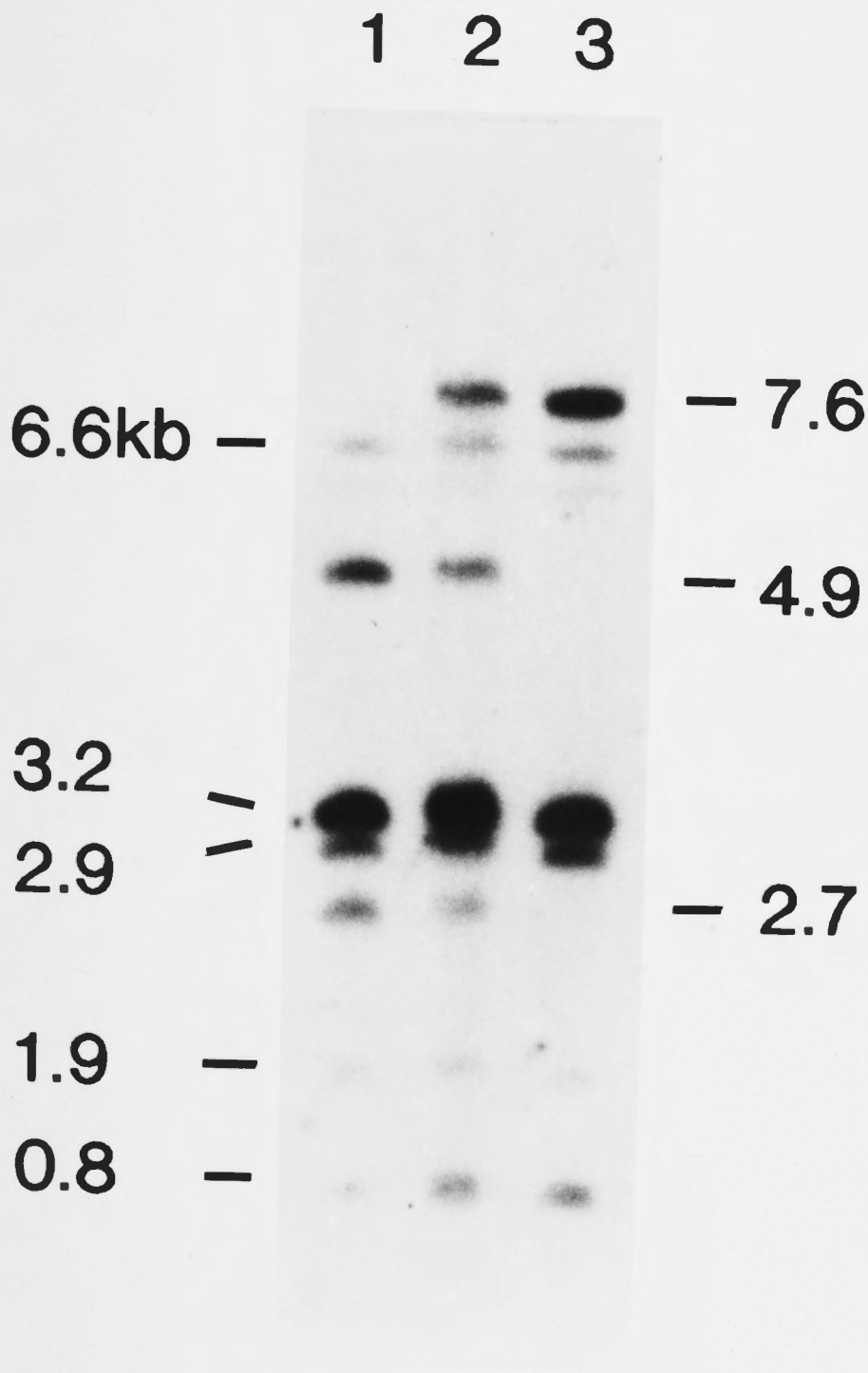
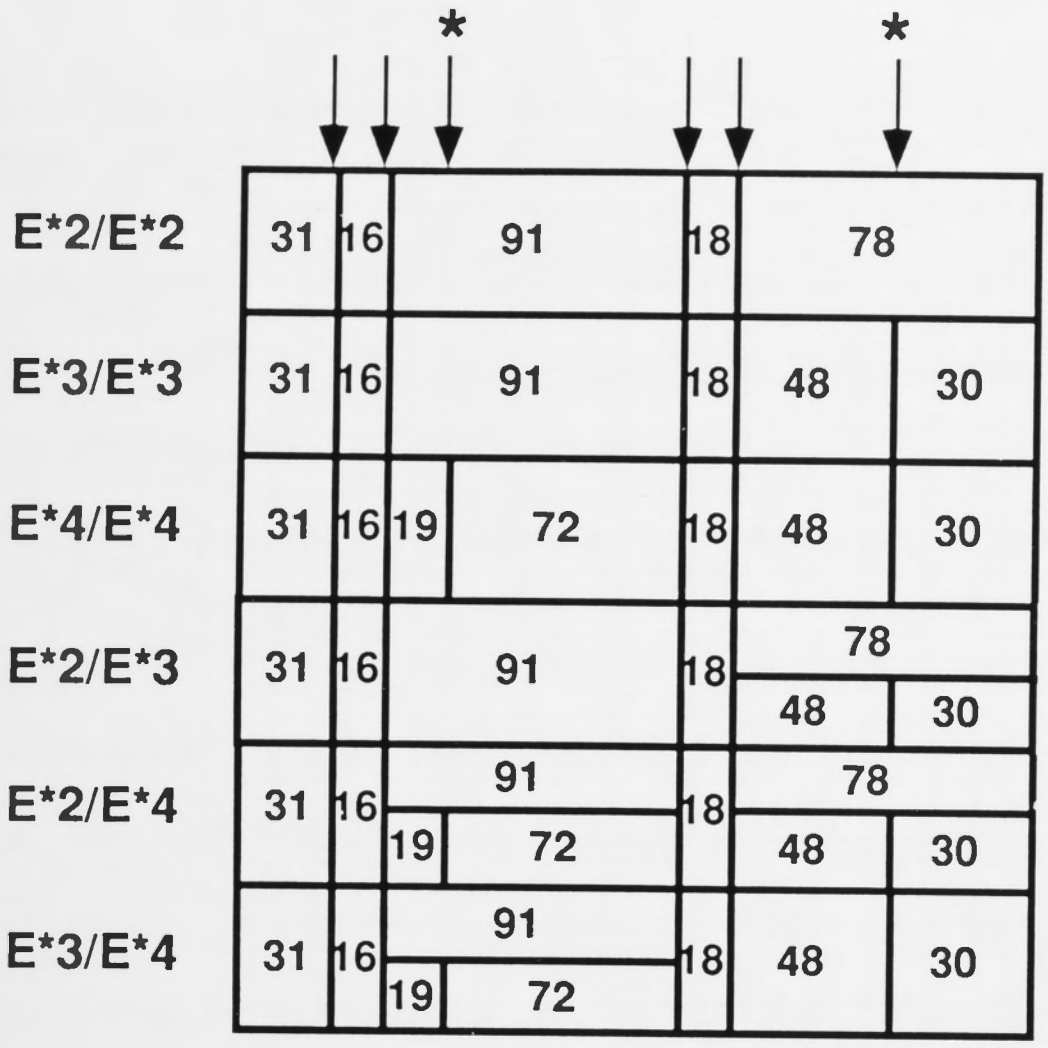
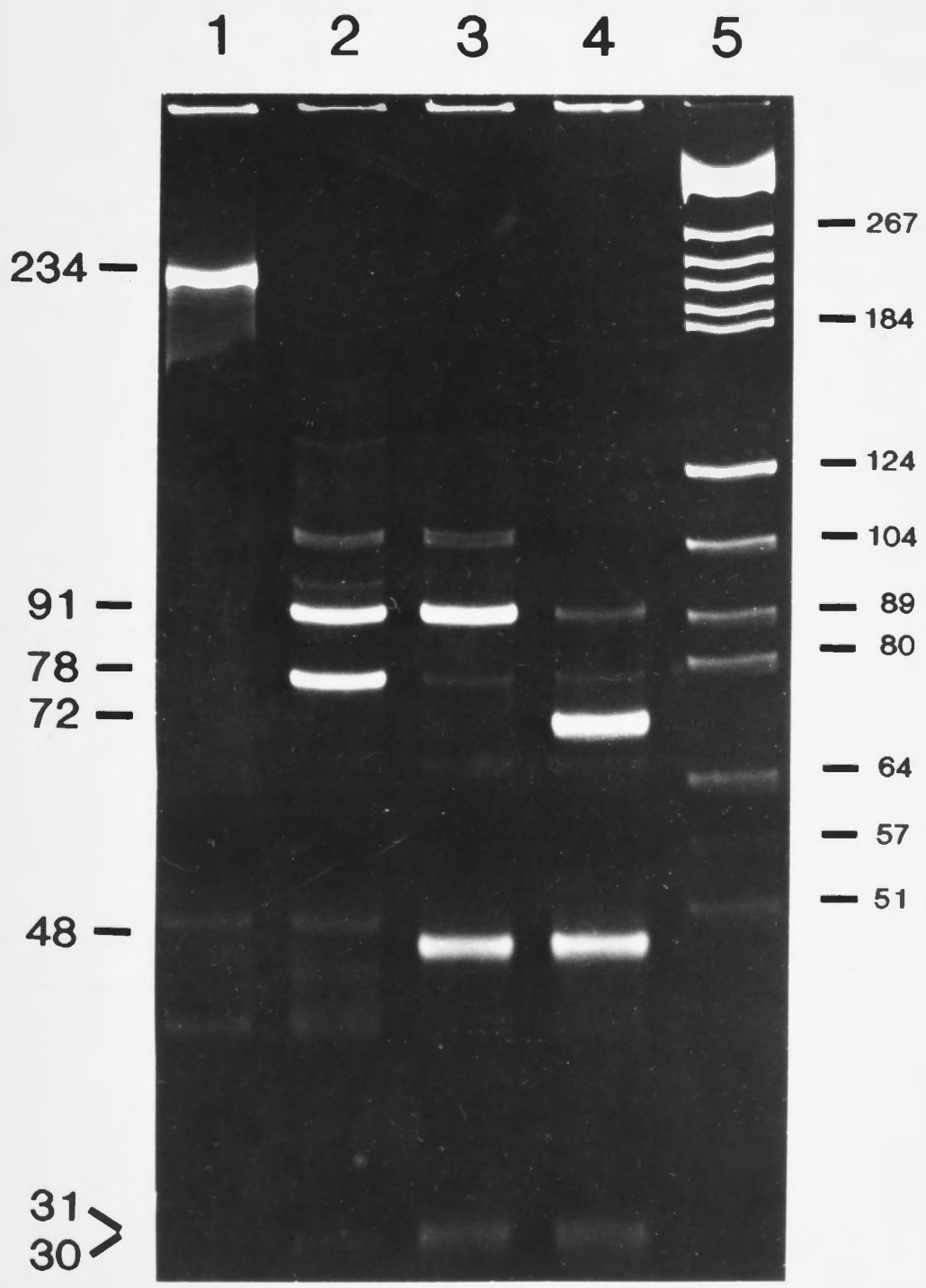


Figure 3.9 Amplified fragment length polymorphism typing of common APOE isoforms. **a** *Cfo*I cleavage map of amplified APOE exon 4 234bp fragment for six the most common APOE phenotypes. Sizes in bp; arrows = cleavage sites; * = polymorphic cleavage sites. **b** Polyacrylamide gel after electrophoresis of uncut 234bp amplification product (lane 1) and *Cfo*I fragments from an APOE*2/APOE*2 homozygote (lane 2), APOE*3/APOE*3 homozygote (lane 3), and APOE*4/APOE*4 homozygote (lane 4). The fragment sizes (in bp) of a DNA standard (*Hae*III digested pBR322) are shown in lane 5.

a


*
*
E*2/E*2
31 16 91 18 78
E*3/E*3
31 16 91 18 48 30
E*4/E*4
31 16 19 72 18 48 30
E*2/E*3
31 16 91 18 78 48 30
E*2/E*4
31 16 91 19 72 18 78 48 30
E*3/E*4
31 16 91 19 72 18 48 30

b


1 2 3 4 5
234
91
78
72
48
31
30
267
184
124
104
89
80
64
57
51

3.4.3 Micronesian lipoprotein genotype distributions and allele frequencies

Reliable typings were obtained for samples of between 74 and 101 for the eight lipoprotein genes under study. The genotype distributions are presented in Table 3.8. All of the distributions were consistent with expections based on Hardy-Weinberg equilibrium.

Frequencies for the presence of any one site for the diallelic systems ranged from a low of 0.230 ± 0.035 in LDLR-P1 to a high of 0.839 ± 0.027 for APOC1-D1 (Table 3.9). Direct-count heterozygosity ranged from 0.233 for APOC1 to 0.564 for APOC2, with an average of 0.395 (SE 0.040). Seven of the eight loci examined were "reasonably informative" (0.25<PIC<0.50), while APOC1 was only "slightly informative" (PIC<0.25), according to the criteria of Botstein *et al.* (1980).

Table 3.8 Micronesian genotype frequencies for APOA1, APOA2, APOC1, APOC2, APOD, APOE, LDLR and LPL.

Locus (Enzyme)	Genotype	Observed		Expected	χ^2	p value†
		N	%	N		
APOA1	M1-M1	13	14.1	15.3	0.97	0.324
(*Msp*I)	M1-M2	49	53.3	44.4		
	M2-M2	30	32.6	32.3		
	Total	92				
APOA2	M1-M1	50	49.5	49.2	0.14	0.709
(*Msp*I)	M1-M2	41	40.6	42.6		
	M2-M2	10	9.9	9.2		
	Total	101				
APOC1	D1-D1	65	72.2	63.4	....	0.231
(*Dra*I)	D1-D2	21	23.3	24.3		
	D2-D2	4	4.5	2.3		
	Total	90				
APOC2	T1-T1	25	24.8	28.4	1.78	0.182
(*Taq*I)	T1-T2	57	56.4	50.3		
	T2-T2	19	18.8	22.3		
	Total	101				
APOD	T1-T1	13	15.5	11.1	0.83	0.364
(*Taq*I)	T1-T2	35	41.7	38.8		
	T2-T2	36	42.8	34.1		
	Total	84				
APOE	E*2/E*2	0	0.0	0.2	....	1.000
(*Cfo*I)	E*2/E*3	7	8.3	6.6		
	E*2/E*4	1	1.2	1.0		
	E*3/E*3	56	66.7	57.5		
	E*3/E*4	20	23.8	17.4		
	E*4/E*4	0	0.0	1.3		
	Total	84				
LDLR	P1-P1	3	4.1	3.9	....	0.747
(*Pvu*II)	P1-P2	28	37.8	26.2		
	P2-P2	43	58.1	43.9		
	Total	74				
LPL	P1-P1	49	66.2	48.7	....	0.713
(*Pvu*II)	P1-P2	22	29.7	22.7		
	P2-P2	3	4.1	2.6		
	Total	74				

† Testing Hardy-Weinberg equilibrium contingency χ^2 with d.f.=1, or exact probability.

Table 3.9 Micronesian allele frequencies for APOA1, APOA2, APOC1, APOC2, APOD, APOE, LDLR and LPL.

Locus (Enzyme)	Allele	Frequency	SE	Hetero-zygosity†	PIC value
APOA1	M1	0.408	0.036	0.533	0.366
(*Msp*I)	M2	0.592	0.036		
	N	92			
APOA2	M1	0.698	0.032	0.406	0.333
(*Msp*I)	M2	0.302	0.032		
	N	101			
APOC1	D1	0.839	0.027	0.233	0.234
(*Dra*I)	D2	0.161	0.027		
	N	90			
APOC2	T1	0.530	0.035	0.564	0.374
(*Taq*I)	T2	0.470	0.035		
	N	101			
APOD	T1	0.363	0.037	0.417	0.356
(*Taq*I)	T2	0.637	0.037		
	N	84			
APOE	E*2	0.048	0.016	0.333	0.273
(*Cfo*I)	E*3	0.827	0.029		
	E*4	0.125	0.026		
	N	84			
LDLR	P1	0.230	0.035	0.378	0.291
(*Pvu*II)	P2	0.770	0.035		
	N	74			
LPL	P1	0.811	0.032	0.297	0.260
(*Pvu*II)	P2	0.189	0.032		
	N	74			

† Heterozygosity calculated by direct counting.

3.4.4 Comparisons with other ethnic groups

The APOA-M1 allele (0.408 ± 0.036, PIC 0.366) was less common, and with a frequency closer to 0.5, more informative, in Micronesians than in either Caucasoids or Africans (Table 3.1). Comparisons with the Caucasoid and African populations with the closest allele frequencies (0.88 ± 0.03, PIC 0.220 and 0.74 ± 0.06, PIC 0.318 respectively, Coleman *et al.*, 1986; Paul *et al.*, 1987)

indicate that the differences in allele frequencies were significant (Caucasoid χ^2=72.24, d.f.=1, p=<0.001; African χ^2=17.24, d.f.=1, p<0.001). The Micronesian APOA1 allele frequencies were not dissimilar from those most recently reported for Japanese (0.44 ± 0.04, Thompson *et al.*, 1988) or from Asian Indians (0.54 ± 0.07, Paul *et al.*, 1987) (Japanese χ^2=0.24, d.f.=1, p=0.627; Asian Indians χ^2=2.24, d.f.=1, p=0.136).

Caucasoid allele frequencies for APOA2-M1 were variable (Table 3.2), yet the lowest of the reported frequencies (0.81 ± 0.03, PIC 0.260, Scott *et al.*, 1985) was significantly higher than that found in Micronesians (0.698 ± 0.032, PIC 0.333) (χ^2=5.71, d.f.=1, p=0.017). The only reported Caucasoid APOC1-D1 frequency of 0.786 ± 0.015 (PIC 0.280, Frossard *et al.*, 1987) was similar to that found in the Micronesian sample (0.839 ± 0.027, PIC 0.234; χ^2=2.21, d.f.=1, p=0.137).

The frequency of the presence of the APOC2 *Taq*I site in Micronesians (0.530 ± 0.035, PIC 0.374) was similar to that in both African and Asian groups (Williams *et al.*, 1985 African 0.48 ± 0.05; West Indian African 0.52 ± 0.05; Asian Indian 0.50 ± 0.05; Chinese 0.58 ± 0.06; Japanese 0.56 ± 0.06; p values for comparisons between 0.489 and 0.958). The frequency of the APOC2-T1 allele was, however, significantly increased over that found in pooled Caucasoid samples (0.41 ± 0.02, PIC 0.367, Table 3.3; χ^2=8.91, d.f.=1, p=0.003).

The frequency of the APOD *Taq*I site was significantly increased, and more informative in Micronesians (0.363 ± 0.037, PIC 0.356), than in Caucasoids (0.18 ± 0.03, PIC 0.252, Drayna *et al.*, 1987*b*; χ^2=15.18, d.f.=1, p<0.001).

Both the APOE*2 and APOE*4 (0.048 ± 0.016 and 0.125 ± 0.026) frequencies fall within the range of those reported for Caucasoid populations (APOE*2 0.039 ± 0.004 to 0.130 ± 0.016; APOE*4 0.116 ± 0.013 to 0.227 ± 0.012; Table 3.4). The APOE*3 frequency (0.827 ± 0.029) was slightly higher than in Caucasoids, but not significantly different (χ^2=1.25, d.f.=1, p=0.263) from the highest reported frequency of 0.788 ± 0.008 (Ordovas *et al.*, 1987).

The frequencies for all three APOE alleles differed from those found in two African populations, and from Chinese and New Guinean samples (African χ^2=19.86, d.f.=2, p<0.001; Chinese χ^2=7.60, d.f.=2, p=0.022; New Guinean χ^2=47.75, d.f.=2, p<0.001; Table 3.4; Davignon *et al.*, 1988; Kamboh *et al.*, 1990). The larger of the Japanese surveys (Eto *et al.*, 1986*a*) and the Mexican sample (Hanis *et al.*, 1991) report APOE allele frequencies which were similar to those found in Micronesians (Japanese χ^2=0.54, d.f.=2, p=0.763; Mexican χ^2=1.28, d.f.=2, p=0.528). Mayans and Amerindians did not possess the APOE*2 allele (Asakawa *et al.*, 1985; Kamboh *et al.*, 1991). Although it is possible that the sporadic occurrence (5%) of APOE*2 in Micronesians is a consequence of ancestral foreign admixture, this seems unlikely because Melanesians from Papua New Guinea have the

highest frequency of APOE*2 yet reported (Kamboh *et al.*, 1990).

The LDLR-P1 frequency of 0.230 ± 0.035 was within the reported range for Caucasoids (0.18 ± 0.02 to 0.402 ± 0.033; Table 3.5). The presence of the LPL *Pvu*II site was more common (χ^2=33.09, d.f.=1, p<0.001) and less informative, in Micronesians (0.811 ± 0.032, PIC 0.260) than in a pooled sample of Caucasoids (0.552 ± 0.017, PIC 0.372; Table 3.6).

3.5 DISCUSSION

3.5.1 Comparisons with other ethnic groups

All seven RFLPs and the APOE polymorphism, chosen on the basis of reported polymorphism in Caucasoids, were found to be informative and polymorphic in Micronesians. This is fortuitous, if not a little surprising, as the same polymorphisms need not necessarily be expected to occur in different ethnic groups. The erythrocyte-type glucose transporter (GLUT1) gene was found to be associated with *Bgl*II RFLPs in Micronesians (Serjeantson *et al.*, 1989), but not in Caucasoids (Shows *et al.*, 1987). Conversely many HLA Class I antigens found in Caucasoids were not present in Micronesians (Serjeantson *et al.*, 1983). Three HLA DPα/DPβ RFLP defined haplotypes, present in Caucasoids were reported absent in Micronesians, whilst two haplotypes, present in Micronesians, were reported absent in Caucasoids (Easteal *et al.*, 1989). There is then, no precedent for expecting equivalent polymorphisms to occur in Micronesians and Caucasoids.

No consistent pattern of difference was found between lipoprotein allele frequencies in Micronesians and other ethnicities. Of the eight loci examined, five had allele frequencies significantly different from Caucasoids. Of these APOA1 and LPL differed most markedly (by 0.472 and 0.259 respectively). APOC2 presented the most consistent allele frequencies across ethnic groups, and although significant, the difference between Micronesians and Caucasoids at this locus was only 0.120. Micronesian and Caucasoid allele frequencies did not differ at the APOC1, APOE or LDLR loci.

The small number of studies in other ethnicities necessitates caution in drawing conclusions, but preliminary comparisons indicate that the Micronesian lipoprotein allele frequencies are dissimilar to those of Caucasoids and Africans and closer to those of Orientals. Similarities existed at four of five loci compared for Orientals, but at only two of the five compared for Africans.

3.5.2 Heterozygosity in Micronesian lipoprotein genes

Nauru supported a population of about 1000 to 1400 in pre-European times (Simpson, 1844). Availability of water on the island was unpredictable, with rainfall known to vary between 13 to 460cm a year (Viviani, 1970). It is therefore likely that the Nauruans were exposed to times of drought and famine, with consequent fluctuations in population size. Flint *et al.* (1989) suggest that such circumstances of potential bottleneck effects, or the maintenance of a

small population size over many generations may result in lowered heterozygosity. However, Maruyama and Fuerst (1985) make the suggestion, based on theoretical observations, that after a bottleneck event, allelic loss may occur more rapidly than loss of genic heterozygosity. The average heterozygosity of 0.395 ± 0.040 in the Micronesian population is not low relative to the maximum heterozygosity for a diallelic system of 0.5.

Polynesians, traditionally existing, at times, under similar conditions as the Nauruans, have been found to have low diversity as revealed by examination of α-globin haplotypes and tandem repetitive highly variable loci (Flint *et al.*, 1989; Hertzberg *et al.*, 1988). The extent of this low diversity had not previously been realised in studies of less polymorphic systems, perhaps because the ability to detect only electrophoretic variation had masked a sizable proportion of the diversity present in a given population. This may also be true of the lipoprotein loci in Micronesians; the RFLP and AFLP data presented here may suggest relatively high heterozygosity, but this can only be verified using more polymorphic markers. Variability, measured at the level of RFLP and AFLP data, has however been maintained at these eight loci in the Micronesian population.

3.5.3 Micronesian lipoprotein genes as candidate markers for disease studies

The prerequisites for using RFLPs and AFLPs as genetic markers are twofold. First these markers must be

polymorphic in the study population. This criterion is met by the eight lipoprotein fragment length polymorphisms under study. Second, given the presence of variability, the information content of the polymorphisms must be such as to maximize the probability of detection of linkage to the disease locus. The polymorphic information content is calculated as the probability that a given offspring of a parent carrying the predisposing allele at the disease locus will allow deduction of the genotype at the marker locus. For any marker this is evaluated by multiplying the frequencies of all possible parental matings by the probability that an offspring will be informative (Botstein *et al.*, 1980).

High PIC values result when the marker locus has a high number of alleles and equitable allele frequencies. Consequently maximum PIC values of 0.375, for a two allele system, and 0.602 for a three allele system, are obtained where $p = q = 0.50$ and $p = q = r = 0.33$ respectively (where p, q, and r are the allele frequencies for alleles 1, 2, and 3 respectively). These PIC values may be as high as those of multiple allele systems where one or two alleles predominate and the remaining alleles are rare.

The average PIC for the seven RFLPs examined was 0.316, or 84.2% (with a range of 62.4% to 99.7%) of the maximum for diallelic systems. The APOE locus had a relatively low PIC value, being only 0.273 or 45% of the maximum for a three allele system. The APOA1, APOA2 and APOD RFLPs had PIC values markedly higher than in Caucasoids (increased by

38.9%, 19.5% and 27.7% respectively). The converse was true for APOC1 and LPL (decreased by 12.3% and 29.9% respectively). Overall, the loci examined, except APOC1, were reasonably informative on the PIC scale of zero to one, and, except APOE, highly informative within the limits set by the number of alleles.

The apolipoprotein and lipoprotein genes are known to be involved in the monogenic determination of hypo- or hyperlipoproteinemias (Falus and Romics, 1988; Assmann *et al.*, 1990; Hata *et al.*, 1990*b*). They are also strongly implicated in the aetiology of polygenic variation in lipoprotein levels (Galton *et al.*, 1990; Pedersen and Berg, 1990). The eight genes examined here therefore meet the criteria for consideration as candidate genes for the lipoproteinemias. As lipoprotein metabolism is altered in diabetes mellitus (Brunzell *et al.*, 1985; Assmann and Schulte, 1988) they can also be regarded as candidate genes for NIDDM.

3.6 CONCLUSIONS

1) The lipoprotein genes APOA1, APOA2, APOC1, APOC2, APOD, APOE, LDLR and LPL have been examined in a Micronesian population and either RFLP or functional allele frequencies estimated. All eight loci were polymorphic. Frequencies for the presence of restriction sites ranged from 0.230 ± 0.035 to 0.839 ± 0.027. APOE allele frequencies were APOE*2 0.048 ± 0.016, APOE*3 0.827 ± 0.029 and APOE*4 0.125 ± 0.00.026.

2) The allele frequencies for Micronesians have been compared with those for other ethnic groups. Micronesian lipoprotein allele frequencies were found to be dissimilar to those of Caucasoids and Africans and closer to those of Orientals.

3) The seven RFLPs had a heterozygosity of 0.395 ± 0.040 and were highly informative, given that they were dimorphic. The APOE polymorphism was found to be reasonably informative. The lipoprotein gene polymorphisms under study meet the criteria required for use as candidate disease markers in NIDDM and the lipoproteinemias.

SECTION 4

ASSOCIATIONS OF MICRONESIAN LIPOPROTEIN GENES WITH LIPOPROTEINEMIA, OBESITY AND DIABETIC STATUS

4.1 AIMS

General aims:- To investigate the genetic contribution of lipoprotein genes to lipoproteinemia, obesity and non-insulin diabetes mellitus in Micronesians.

Specific aims:-

1) To investigate the influence of lipoprotein gene RFLP frequencies and apolipoprotein E allele frequencies on plasma cholesterol and triglyceride concentrations in a Micronesian population.

2) To investigate the influence of lipoprotein gene RFLP frequencies and apolipoprotein E allele frequencies on obesity, as measured by body mass index, and on degree of central adiposity, as measured by waist-to-hip ratio in a Micronesian population.

3) To investigate the influence of lipoprotein gene RFLP frequencies and apolipoprotein E allele frequencies on non-insulin dependent diabetes mellitus in a Micronesian population.

4.2 INTRODUCTION

4.2.1 Lipoprotein genes in lipoproteinemia - population studies

Several associations have been documented between apolipoprotein and lipoprotein RFLPs and hyperlipoproteinemias or atherosclerosis (Section 1.3.3).

The following discussion will be restricted to include only reported investigations of the RFLPs and the APOE AFLPs examined in this study.

RFLP variation at the chromosome 11 apolipoprotein gene cluster has been shown to be associated with both angiographically defined atherosclerosis and the occurrence of myocardial infarct (Section 1.3.3 and Galton, 1990). The situation regarding associations of variation at this locus with hyperlipoproteinemia is less clear. The APOA1 *Msp*I site has been found in weak association with hyperlipoproteinemia in Caucasoids (Ferns and Galton, 1986*a*) but not Japanese (Rees *et al.*, 1986). Freeman *et al.* (1990) found a significant difference in plasma LCAT activity between APOA1-M1 homozygotes and heterozygotes. As LCAT is responsible for the formation of nearly all plasma cholesteryl esters in normal individuals (Glomset, 1968), differences in its activity are expected to affect plasma cholesterol levels.

The *Msp*I site at the APOA2 locus has been the subject of five published studies concerning hyperlipoproteinemia. APOA2-M1 was significantly associated with lowered HDL levels and increased triglycerides in Caucasoids (Deeb *et al.*, 1986). This finding was in agreement with a previous report of altered HDL composition (apolipoprotein A1/apolipoprotein A2 ratio) in APOA2-M2 homozygotes (Scott *et al.*, 1985) and with the results of Ferns *et al.* (1986*b*) who found the APOA2-M1 allele to be more common in hypertriglyceridemics. However, an absence of association

between the *Msp*I site and a range of plasma lipoprotein levels, including triglycerides and HDL-cholesterol has recently contradicted these earlier findings (Rajput-Williams *et al.*, 1989).

Whilst the APOC2 *Taq*I site has been used to establish linkage to apolipoprotein C2 deficiency phenotypes in type I hyperlipoproteinemia (Humphries *et al.*, 1984), no associations have been reported between allele frequencies and hypertriglyceridemia. There is a reported absence of association between this *Taq*I RFLP and types IIa, IIb, III, IV and V hyperlipoproteinemia (Humphries *et al.*, 1983). The APOC2-T1 allele was, however, found to be associated with increased cholesterol levels (Deeb *et al.*, 1986).

The apolipoprotein E polymorphism has been well studied. In addition to being associated with differences in plasma triglyceride levels this polymorphism also explains a consistent, but small proportion of inter-individual variability in plasma cholesterol levels (Section 1.3.2.6; Robertson and Cumming, 1985; Sing and Davignon, 1985; Boerwinkle and Utermann, 1988; Davignon *et al.*, 1988).

In general, APOE*2 has been associated with lower plasma cholesterol concentrations, APOE*4 with the higher, and APOE*3 with intermediate levels. APOE*2 has also been shown to be associated with higher levels of plasma triglycerides (Assmann *et al.*, 1984; Sing and Davignon, 1985; Ehnholm *et al.*, 1986). One exception to this pattern of association is seen in type III hyperlipoproteinemia, where 90% of patients possess the APOE*2/APOE*2 phenotype (Utermann *et*

al., 1979; Breslow *et al*., 1982a). Type III hyperlipoproteinemia is characterized by elevated levels of plasma cholesterol and triglycerides, and premature heart disease (Refer Table 1.3). As the APOE*2/APOE*2 phenotype is present in 1 per 100 Caucasoids, and type III hyperlipoproteinemia has a prevalence of only 1 per 10,000, it is generally accepted that other factors are required before the disease is expressed (Utermann, 1985). The situation is further complicated as Ghiselli *et al*. (1982) and Utermann *et al*., (1984) report an association of APOE*4 with hypertriglyceridemia (types IIb and V hyperlipoproteinemia).

The *Pvu*II site in LDLR has been used alone, or in concert with other polymorphic sites, to establish disease genotypes or haplotypes in familial hypercholesterolemia (Kotze *et al*., 1987; Daga *et al*., 1988; Kotze *et al*., 1989; Schuster *et al*., 1989). The presence of the site has also been found to be associated with decreased total cholesterol (Brink *et al*., 1986; Humphries *et al*., 1991) and LDL-cholesterol (Humphries *et al*., 1991), and it has recently been shown that the presence of the site eliminates the effect of APOE*4 on cholesterol levels (Berg, 1990). The *Pvu*II site is thought to account for 3% of the sample variance in total cholesterol levels (Schuster *et al*., 1990) and 9.6% of LDL-cholesterol variability (Humphries *et al*., 1991). Individuals homozygous for the presence of the *Pvu*II site in LPL have lowered triglycerides levels (Chamberlain *et al*., 1989).

The APOA1 *Msp*I, APOA2 *Msp*I, APOC1 *Dra*I and APOC2 *Taq*I RFLPs have recently been examined in a study analyzing associations of apolipoprotein RFLPs with lipid abnormalities. Myklebost *et al.* (1990) report a lack of association between RFLP variation at these sites and variation in plasma lipoprotein levels.

4.2.2 Lipoprotein genes in non-insulin dependent diabetes mellitus - population studies

Relatively few studies have examined associations of lipoprotein genes with NIDDM. In a small study of 45 Caucasoid hypertriglyceridemics, 15 of whom had NIDDM, an association was found between the *Sst*I site in the APOA1-C3-A4 cluster and triglyceridemia but not NIDDM (Jowett *et al.*, 1984). The frequency of the S2 allele at this site was, however, increased in patients with NIDDM and concurrent coronary heart disease (CHD) (14%, N=47) when compared with non-diabetic controls (2%, N=35) (Trembath *et al.*, 1987). Normotriglyceridemic diabetics without CHD did not show the increase in S2 frequency which was observed in those with CHD. It was concluded that there was no association between diabetes alone and the *Sst*I site.

Within the same apolipoprotein cluster a 2.5kb *Eco*RI fragment was significantly associated with NIDDM in a Polish population. In a sample of 100 patients with diabetes, five were homozygous for the fragment and eight heterozygous, compared with only one homozygote and one heterozygote in a sample of 100 controls. Both controls

with the 2.5kb fragment had a family history of diabetes (Buraczynska *et al.*, 1985).

The APOA2, APOB and APOA1-C3-A4 loci have been examined in Chinese Americans with NIDDM (Xiang *et al.*, 1989). The APOB *Xba*I RFLP appeared to contribute to the development of NIDDM in individuals of lean/normal weight. Significant differences in the distribution of the *Msp*I alleles alone, and of *Msp*I/*Pst*I/*Sst*I haplotypes at the APOA1-C3-A4 locus were reported between overweight diabetic and non-diabetic groups. Most of the difference was observed at the *Msp*I site, with the M1 allele having an increased frequency in overweight patients with diabetes.

The common APOE genetic variants are known to contribute to normal plasma lipoprotein variation (Boerwinkle and Utermann, 1988), with the APOE*2 and APOE*4 alleles being more frequent in hypertriglyceridemia and hypercholesterolemia respectively (Sections 1.3.2.6 and 4.2.1; Assmann *et al.*, 1984; Utermann *et al.*, 1984). In addition, the VLDL of NIDDM patients contains more apolipoprotein E than does that of hyperlipoproteinemics (Black *et al.*, 1990). The APOE*2 variant in particular is implicated *a priori* in the pathogenesis of NIDDM due to the presence of hypertriglyceridemia in NIDDM.

The phenotype distributions of APOE were similar in NIDDM and control groups in both Japanese and Caucasoid populations (Eto *et al.*, 1986*b*; Imari *et al.*, 1988; Snowden *et al.*, 1991). However, Vogelberg and Maucy (1988) have detected an increase in APOE*2 containing phenotypes in

NIDDM, due to an increase in the frequency of APOE*2 homozygotes in the diabetics.

When the degree of lipoproteinemia, but not obesity, in NIDDM patients is considered associations previously reported for non-NIDDM hyperlipoproteinemic individuals were maintained. Both APOE*4 containing phenotypes and APOE*4 allele frequencies were increased in Japanese hypercholesterolemics (type IIa and IIb) with NIDDM (Eto *et al.*, 1987) and the APOE*2 allele was found more often in Japanese hypertriglyceridemic patients with NIDDM than in normolipidemics (Imari *et al.*, 1988). In a population of unstated ethnicity, the APOE*2/APOE*3 and APOE*3/APOE*4 phenotypes were more common in hyperlipoproteinemics with NIDDM than normolipidemics. In this population the APOE*2 allele was associated with hypertriglyceridemia, and the APOE*4 with mixed lipoproteinemias and hypercholesterolemia (Parhofer *et al.*, 1990).

4.2.3 Lipoprotein genes, obesity and non-insulin dependent diabetes mellitus

Obesity has been recognised as an important risk factor for NIDDM (WHO Expert Committee, 1980). It is now thought that obesity itself is less of a cause of NIDDM than the duration of the obesity, and the degree of associated central adiposity. These factors may act to precipitate the onset of NIDDM in genetically susceptible individuals (Zimmet *et al.*, 1989). Consequently it may prove informative to investigate patterns of obesity (as measured by BMI) and degree of central adiposity (as measured by

WHR) in relation to genetic variation in the lipoprotein genes since they are candidate genes for NIDDM. This approach has proven fruitful in the case of APOB and the APOA1-C3-A4 loci (Section 4.2.2) which showed associations with NIDDM, mediated by the degree of obesity.

The relationship between APOE alleles and obesity or central adiposity may be of particular interest as triglycerides are raised in both APOE*2 carriers and in NIDDM patients. Two reports indicate that obese subjects with APOE*2 or APOE*4 were not significantly more susceptible to hyperlipoproteinemia than obese subjects with the common APOE*3/APOE*3 phenotype (Eto *et al.*, 1989; Gueguen *et al.*, 1989), although Gueguen *et al.*, (1989) did find an interactive effect between APOE*4 and weight change on triglyceride levels. In support of the latter finding Fumeron *et al.* (1988) claimed that the APOE*4 carrying phenotypes were indeed more often associated with hypertriglyceridemia in obesity than were the APOE*3 phenotypes. The effect of APOE*4 in increasing triglycerides in obese subjects in general was also present in obese subjects with NIDDM (Eto *et al.*, 1986*b*, 1987).

This finding appears in conflict with the general rule of associations in the general population. The association which has been cited between APOE*4 and hypertriglyceridemia in the general population (Ghiselli *et al.*, 1982) may be artifactual since this study claims the association with type V hyperlipoproteinemia without excluding subjects with obesity, which is frequent in this

disorder. The relationships between NIDDM, apolipoprotein E phenotypes, hyperlipoproteinemia and obesity remain unclear. However, Fumeron *et al.* (1988) suggest that in the non-obese population APOE*4 is associated with hypercholesterolemia, and that this association is augmented in NIDDM. In the obese population they propose that, even in the presence of NIDDM, the main association is between APOE*4 and hypertriglyceridemia.

4.3 MATERIALS AND METHODS

4.3.1 Sample selection

The series of samples referred to in Section 3.3.1 were used in the analysis of allele and genotype distribution frequencies in relation to plasma cholesterol concentrations (CHO), plasma triglyceride concentrations (TG), body mass index (BMI) and waist-to-hip ratio (WHR). Measurements of CHO, TG, BMI and WHR were obtained as described in Section 2.1.3.

Further individuals with NIDDM were selected from the 1987 Nauru Survey blood sample collections with the aim of increasing the proportion of NIDDM patients in the series. Where additional NIDDM individuals were found to have a non-NIDDM first degree relative represented in the original series this relative was excluded. This extended series was used in the analysis of allele and genotype distribution in relation to NIDDM.

4.3.2 APOA1, APOA2, APOC1, APOC2, APOD, LDLR and LPL RFLP typing

Genomic DNA was prepared from the samples, digested with *Dra*I, *Msp*I, *Pvu*II and *Taq*I and transferred to nylon membranes as described in Chapter 2. The cut DNA was hybridized to plasmids containing the cDNA probes discussed in Section 2.2.4 (refer also Table 3.7).

4.3.3 APOE AFLP typing

A 234bp fragment of APOE exon 4 was amplified from each sample and digested with *Cfo*I according to the methods described in Chapter 2.

4.3.4 Statistical methods

Sample allele frequencies were calculated and contingency testing performed as described in Section 3.3.5. Equality of variance was tested before *t*-tests were performed and either separate variance or pooled variance estimators used accordingly (Snedecor and Cochran, 1980). Pearson's correlation coefficients are reported at a significance level of $p \leq 0.05$.

Multiple regression and loglinear modelling were carried out in a stepwise manner, with the significance of removal of established independent variables being tested with the entry of each new independent variable. The degree of impact of each independent variable on the dependent variable is reported as β, the standardized regression coefficient. β represents the effect that one standard deviation difference in the independent variable would have

on the standardized scores of the dependent variable. The explanatory power of each new independent variable is reported as the change in R^2 (ChR^2) and is the proportion of variance in the dependent variable explained by variability in the independent variable. The level of significance for all tests carried out in association with the multiple regression and loglinear analyses, unless otherwise stated, was 5%. The reported significance test for entry of an independent variable into the regression is the change in F (FCh), the ratio of the change in sum of squares over the residual sum of squares, with associated probability SigFCh. As the distribution of plasma triglycerides was skewed, triglyceride values were transformed by the function:- transformed TG (tranTG) = $\log_{10}(TG + 1)$ before entry into either multiple regression or loglinear analyses.

Generalized linear modelling, with a binomial error term, was used to model the effect of continuous and binary independent variables on the binary dependent variable of diabetic status (Healy, 1988). In this situation each individual *x* was treated as a group of size one, with either zero or one successes in the group relating to the presence or absence of diabetes. The change in deviance (ChDev) upon the addition or removal of dependent variables in this case is distributed as a χ^2 (with associated degrees of freedom). An estimate of the proportion of variation (% Var) in the dependent variable due to the addition of a new independent variable is given by the associated change in deviance over the total deviance.

Because large sample approximations break down when the sample at each set of *x*-values is only of size one, the residual deviance could not be used as a measure of the goodness of fit of the model.

4.3.5 Sample description

The total sample of 119 described in Section 3.3.2 was used in the analysis of CHO, TG, BMI and WHR and is referred to as the random series. Three individuals were of unknown diabetic status, while one had unknown CHO and TG, and two had unknown WHR. These individuals were not excluded from the series. Descriptives for this series, as a whole, by sex, and subgrouped into high and low categories for each of the four test variables are presented in Tables 4.1 and 4.2. The top quartile of the sample for each variable was assigned to the high group, whilst those below the top quartile were assigned to the low group. The cut off values for the sample were as follows:- plasma cholesterol 5.86 mM; plasma triglycerides 1.59 mM; body mass index 39.20 kg/m^2; waist-to-hip ratio 0.91.

A total sample of 140 was obtained in the extended NIDDM series. Individuals with missing values for CHO, TG, BMI or WHR were not excluded from the series. Descriptive statistics for this series are presented in Tables 4.1 and 4.2.

4.3.6 Within sample correlations

A number of correlations existed within the data set which affect the interpretation of results as presented in the

following sections. Plasma CHO and TG levels were positively correlated ($p \leq 0.05$) in the random series ($r=0.504$), in both females ($r=0.565$) and males ($r=0.428$) and in the control sample ($r=0.538$), but not the NIDDM sample ($r=0.046$, $p=0.360$). Examination of partial correlation coefficients revealed that the relationship between CHO levels and TG levels in the random series was retained after controlling for age (partial $r=0.530$), sex (partial $r=0.531$), or diabetic status (partial $r=0.516$).

There were significant, but relatively minor positive correlations, between both plasma CHO level and diabetic status ($r=0.164$), and plasma TG level and diabetic status ($r=0.147$). Both of these associations held after controlling for sex (partial $r=0.168$ and partial $r=0.152$ respectively), and the CHO and diabetic status association was maintained after controlling for age (partial $r=0.153$). The correlation between TG levels and diabetic status was, however, lost after controlling for age (partial $r=0.117$, $p=0.111$).

Age and diabetic status were correlated ($r=0.355$) regardless of sex (partial $r=0.390$). Another important association was seen between WHR and sex ($r=0.447$), with higher waist-to-hip ratios being found in men. This correlation was retained even after controlling for age (partial $r=0.473$) and diabetic status (partial $r=0.483$). WHR was also positively related to plasma TG levels ($r=0.277$) regardless of age (partial $r=0.295$), sex (partial $r=0.264$), or diabetes (partial $r=0.260$).

Several secondary correlations were also evident in the data set. For example the positive correlation between WHR and age (r=0.310), whilst maintained across the sexes (partial r=0.318), lost substantial magnitude after controlling for the presence of diabetes (partial r=0.213). The same pattern was true for the correlation of high WHR with presence of diabetes (r=0.301), which was strengthened marginally after controlling for sex (partial r=0.338), but weakened by controlling for age (partial r=0.219). These results would indicate that the primary association within the triad of diabetic status, WHR and age was between age and diabetic status.

Table 4.1 Micronesian series sample descriptives:- age, sex, diabetic status.

		Age (years)			Female		Male		NIDDM		non-NIDDM	
		$\bar{x}$	SE	N	N	%	N	%	N	%	N	%
Random series[†]		40.59	1.25	119	67	56.3	52	43.7	36	31.0	80	69.0
Sex	female	39.55	1.64	67	..	....	..	....	19	52.8	45	56.2
	male	41.92	1.93	52	..	....	..	....	17	47.2	35	43.8
CHO	low[§]	39.38	1.34	90	50	74.6	40	78.4	23[#]	63.9	66	83.5
	high	43.32	2.82	28	17	25.4	11	21.6	13	36.1	13	16.5
TG	low	40.19	1.44	89	56[#]	83.6	33	64.7	24	66.7	62	78.5
	high	40.69	2.37	29	11	16.4	18	35.3	12	33.3	17	21.5
BMI	low	41.32	1.48	90	49	73.1	41	78.8	26	72.2	62	77.5
	high	38.31	2.23	29	18	26.9	11	21.2	10	27.8	18	22.5
WHR	low	38.01[*]	1.31	88	54[#]	83.1	34	65.4	24	66.7	62	79.5
	high	48.14	2.68	29	11	16.9	18	34.6	12	33.3	16	20.5
NIDDM series		42.11	1.17	140	75	53.6	65	45.4	64	45.7	76	54.3
Sex	female	41.24	1.57	75	..	....	..	....	32	50.0	43	56.6
	male	43.12	1.76	65	..	....	..	....	32	50.0	33	43.4
NIDDM	yes	47.05[*]	1.63	64	..	....	..	....	..	....	..	....
	no	37.96	1.52	76	..	....	..	....	..	....	..	....

† Within the random series status was unknown for NIDDM n=3, CHO n=1, TG n=1 and WHR n=2.
§ The top quartile was classified as 'high', the remainder as 'low'.
# Two by two contingency χ^2, d.f.=1, $0.01 \leq p \leq 0.05$.
* Two tailed *t*-test, $p \leq 0.01$.

Table 4.2 Micronesian series sample descriptives:- plasma cholesterol and triglyceride levels, body mass index and waist-to-hip ratio.

		CHO (mM)			TG (mM)			BMI (kg/m^2)			WHR		
		$\bar{x}$	SE	N	$\bar{x}$	SE	N	$\bar{x}$	SE	N	$\bar{x}$	SE	N
Random series[†]		5.23	0.10	118	1.37	0.09	118	34.65	0.70	119	0.86	0.01	117
Sex	female	5.25	0.15	67	1.26	0.12	67	35.14	0.98	67	0.84*	0.01	65
	male	5.20	0.14	51	1.51	0.14	51	34.02	1.00	52	0.89	0.01	52
CHO	low[§]	4.75*	0.08	90	1.16*	0.07	90	34.57	0.76	90	0.86	0.01	88
	high	6.76	0.14	28	2.03	0.27	28	35.06	1.73	28	0.88	0.01	28
TG	low	4.97*	0.11	89	0.95*	0.03	89	34.39	0.87	89	0.85	0.01	88
	high	6.00	0.19	29	2.66	0.22	29	35.58	1.09	29	0.89	0.01	28
BMI	low	5.26	0.12	89	1.45*	0.11	89	31.19*	0.49	90	0.86	0.01	88
	high	5.12	0.19	29	1.12	0.10	29	45.39	0.85	29	0.87	0.01	29
WHR	low	5.19	0.11	88	1.21	0.08	88	34.94	0.85	88	0.84*	0.01	88
	high	5.31	0.27	28	1.83	0.27	28	33.94	1.31	29	0.94	0.01	29
NIDDM[††] *series*		5.29	0.10	139	1.51	0.12	139	34.96	0.63	138	0.87	0.01	137
Sex	female	5.19	0.14	75	1.28	0.11	75	35.76	0.87	75	0.84*	0.01	74
	male	5.41	0.14	64	1.79	0.22	64	34.01	0.89	63	0.90	0.01	63
NIDDM	yes	5.59*	0.15	64	1.83*	0.23	64	35.17	1.04	62	0.90*	0.01	62
	no	5.04	0.12	75	1.24	0.09	75	34.80	0.77	76	0.85	0.01	75

† Within the random series status was unknown for CHO n=1, TG n=1 and WHR n=2.
†† Within the NIDDM series status was unknown for CHO n=1, TG n=1, BMI n=2 and WHR n=3.
§ The top quartile was classified as 'high', the remainder as 'low'.
* Two tailed *t*-test, $p \leq 0.01$.

4.4 RESULTS

4.4.1 Micronesian lipoprotein genes in lipoproteinemia and obesity - population study

4.4.1.1 Allele frequencies and genotype distributions in low and high plasma cholesterol groups

Genotype frequencies and allele frequencies in low and high CHO groups for the seven RFLPs and the APOE polymorphism are presented in Tables 4.3 and 4.4 respectively. Sample sizes ranged from 55 to 76 for the low CHO group and 19 to 26 for the high CHO group. All of the distributions were consistent with expectations under Hardy-Weinberg equilibrium.

There was a significant difference between the allele counts for APOD in the low and high groups with the frequency of APOD-T2 increasing from 0.586 ± 0.044 in the low CHO group to 0.800 ± 0.063 in the high CHO group (Table 4.4). There was a significant absence of APOD-T1 homozygotes from the high CHO group, whilst these homozygotes comprised over 20% of the low CHO group (Fisher's exact test p=0.032). Other comparisons indicated that the allele frequencies for the remaining sites were similar in low and high CHO groups.

Table 4.3 Micronesian lipoprotein genotype frequencies in low and high cholesterol groups.

Locus (Enzyme)	Genotype	Low			High		
		N	%	p†	N	%	p†
APOA1	M1-M1	9	13.0	0.426	4	17.4	1.000
(*Msp*I)	M1-M2	37	53.6		12	52.2	
	M2-M2	23	33.4		7	30.4	
	Total	69			23		
APOA2	M1-M1	37	49.3	0.502	13	52.0	0.276
(*Msp*I)	M1-M2	29	38.7		12	48.0	
	M2-M2	9	12.0		0	0.0	
	Total	75			25		
APOC1	D1-D1	45	70.3	0.101	20	76.9	1.000
(*Dra*I)	D1-D2	15	23.4		6	23.1	
	D2-D2	4	6.3		0	0.0	
	Total	64			26		
APOC2	T1-T1	18	23.7	0.321	7	29.2	1.000
(*Taq*I)	T1-T2	43	56.6		13	54.2	
	T2-T2	15	19.7		4	16.6	
	Total	76			24		
APOD	T1-T1	13	20.3	0.394	0	0.0	0.548
(*Taq*I)	T1-T2	27	42.2		8	40.0	
	T2-T2	24	37.5		12	60.0	
	Total	64			20		
APOE§	E*2/E*2	0	0.0	1.000	0	0.0	1.000
(*Cfo*I)	E*2/E*3	5	7.9		2	9.5	
	E*2/E*4	1	1.6		0	0.0	
	E*3/E*3	43	68.3		13	61.9	
	E*3/E*4	14	22.2		6	28.6	
	E*4/E*4	0	0.0		0	0.0	
	Total	63			21		
LDLR	P1-P1	3	5.4	1.000	0	0.0	0.538
(*Pvu*II)	P1-P2	20	36.4		8	42.1	
	P2-P2	32	58.2		11	57.9	
	Total	55			19		
LPL	P1-P1	35	63.7	1.000	14	73.7	0.371
(*Pvu*II)	P1-P2	18	32.7		4	21.0	
	P2-P2	2	3.6		1	5.3	
	Total	55			19		

† Testing Hardy-Weinberg equilibrium, χ^2 p value, d.f.=1, or exact probability.
§ APOE genotypes pooled into three classes for testing:- common homozygotes; common/rare heterozygotes; rare homozygotes and other heterozygotes.

Table 4.4 Micronesian lipoprotein allele frequencies in low and high cholesterol groups.

Locus (Enzyme)	Allele	Low		High		$\chi^{2\dagger}$	p value
		Freq.	SE	Freq.	SE		
APOA1	M1	0.399	0.042	0.435	0.073	0.19	0.665
(*Msp*I)	M2	0.601	0.042	0.565	0.073		
	N	69		23			
APOA2	M1	0.687	0.038	0.760	0.060	0.97	0.325
(*Msp*I)	M2	0.313	0.038	0.240	0.060		
	N	75		25			
APOC1	D1	0.820	0.034	0.885	0.044	1.13	0.288
(*Dra*I)	D2	0.180	0.034	0.155	0.044		
	N	64		26			
APOC2	T1	0.520	0.041	0.563	0.072	0.27	0.605
(*Taq*I)	T2	0.480	0.041	0.438	0.072		
	N	76		24			
APOD	T1	0.414	0.044	0.200	0.063	6.04	0.014
(*Taq*I)	T2	0.586	0.044	0.800	0.063		
	N	64		20			
APOE	E*2	0.048	0.019	0.048	0.033	§	
(*Cfo*I)	E*3	0.833	0.033	0.810	0.061		
	E*4	0.119	0.029	0.143	0.054		
	N	63		21			
LDLR	P1	0.236	0.040	0.211	0.066	0.11	0.744
(*Pvu*II)	P2	0.764	0.040	0.789	0.066		
	N	55		19			
LPL	P1	0.800	0.038	0.842	0.059	0.33	0.568
(*Pvu*II)	P2	0.200	0.038	0.158	0.059		
	N	55		19			

† Contingency χ^2 of allele counts, d.f.=1.
§ Pooled allele count tests for APOE:- E*2 vs E*3 and E*4, p=0.681; E*3 vs E*2 and E*4, p=0.906; E*4 vs E*2 and E*3, p=0.893.

4.4.1.2 Allele frequencies and genotype distributions in low and high plasma triglyceride groups

The distributions of genotypes in high (sample sizes from 54 to 74) and low (sample sizes from 20 to 26) TG groups were consistent with Hardy-Weinberg expectations (Table 4.5). There was a similarity in allele frequencies between the two groups in the majority of sites examined (Table 4.6). However, there was a tendency towards an increase in the frequency of the APOC2-T1 allele in the high TG group, with the frequency of APOC2-T1 homozygotes rising from 21.6% in low TG to 34.6% in high TG. A corresponding fall in the frequency of APOC1-T2 homozygotes was seen from 21.6% in the low TG group to 11.5% in the high TG group. The differences in homozygote frequencies between the two groups were not significant (for APOC2-T1/APOC2-T1, χ^2=1.11, d.f.=1, p=0.292; for APOC2-T2/APOC2-T2, Fisher's exact test p=0.385).

Table 4.5 Micronesian lipoprotein genotype frequencies in low and high triglyceride groups.

Locus (Enzyme)	Genotype	Low			High		
		N	%	p†	N	%	p†
APOA1 (*Msp*I)	M1-M1	10	14.9	0.504	3	12.0	1.000
	M1-M2	36	53.7		13	52.0	
	M2-M2	21	31.4		9	36.0	
	Total	67			25		
APOA2 (*Msp*I)	M1-M1	36	48.6	1.000	14	53.8	1.000
	M1-M2	31	41.9		10	38.5	
	M2-M2	7	9.5		2	7.7	
	Total	74			26		
APOC1 (*Dra*I)	D1-D1	48	75.0	0.595	17	65.4	0.288
	D1-D2	14	21.9		7	26.9	
	D2-D2	2	3.1		2	7.7	
	Total	64			26		
APOC2 (*Taq*I)	T1-T1	16	21.6	0.322	9	34.6	0.686
	T1-T2	42	56.8		14	53.9	
	T2-T2	16	21.6		3	11.5	
	Total	74			26		
APOD (*Taq*I)	T1-T1	9	15.0	0.355	4	16.7	1.000
	T1-T2	23	38.3		12	50.0	
	T2-T2	28	46.7		8	33.3	
	Total	60			24		
APOE§ (*Cfo*I)	E*2/E*2	0	0.0	0.336	0	0.0	1.000
	E*2/E*3	5	7.9		2	9.5	
	E*2/E*4	0	0.0		1	4.8	
	E*3/E*3	43	68.3		13	61.9	
	E*3/E*4	15	23.8		5	23.8	
	E*4/E*4	0	0.0		0	0.0	
	Total	63			21		
LDLR (*Pvu*II)	P1-P1	3	5.6	1.000	0	0.0	0.548
	P1-P2	20	37.0		8	40.0	
	P2-P2	31	57.4		12	60.0	
	Total	54			20		
LPL (*Pvu*II)	P1-P1	35	64.8	0.671	14	70.0	1.000
	P1-P2	16	29.6		6	30.0	
	P2-P2	3	5.6		0	0.0	
	Total	54			20		

† Testing Hardy-Weinberg equilibrium, χ^2 p value, d.f.=1, or exact probability.
§ APOE genotypes pooled into three classes for testing:- common homozygotes; common/rare heterozygotes; rare homozygotes and other heterozygotes.

Table 4.6 Micronesian lipoprotein allele frequencies in low and high triglyceride groups.

Locus (Enzyme)	Allele	Low		High		$\chi^{2\dagger}$	p value
		Freq.	SE	Freq.	SE		
APOA1	M1	0.418	0.043	0.380	0.069	0.22	0.642
(*Msp*I)	M2	0.582	0.043	0.620	0.069		
	N	67		25			
APOA2	M1	0.696	0.038	0.731	0.061	0.22	0.636
(*Msp*I)	M2	0.304	0.038	0.269	0.061		
	N	74		26			
APOC1	D1	0.859	0.031	0.788	0.057	1.37	0.241
(*Dra*I)	D2	0.141	0.031	0.212	0.057		
	N	64		26			
APOC2	T1	0.500	0.041	0.615	0.067	2.06	0.152
(*Taq*I)	T2	0.500	0.041	0.385	0.067		
	N	74		26			
APOD	T1	0.342	0.043	0.417	0.071	0.83	0.361
(*Taq*I)	T2	0.658	0.043	0.583	0.071		
	N	60		24			
APOE	E*2	0.040	0.017	0.071	0.040	§	
(*Cfo*I)	E*3	0.841	0.033	0.786	0.063		
	E*4	0.119	0.029	0.143	0.054		
	N	63		21			
LDLR	P1	0.241	0.041	0.200	0.063	0.27	0.601
(*Pvu*II)	P2	0.759	0.041	0.800	0.063		
	N	54		20			
LPL	P1	0.796	0.039	0.850	0.056	0.55	0.459
(*Pvu*II)	P2	0.204	0.039	0.150	0.056		
	N	54		20			

† Contingency χ^2 of allele counts, d.f.=1, excluding APOE, d.f.=2.

§ Pooled allele count tests for APOE:- E*2 vs E*3 and E*4, p=0.681; E*3 vs E*2 and E*4, p=0.556; E*4 vs E*2 and E*3, p=0.893.

4.4.1.3 Allele frequencies and genotype distributions in low and high body mass index groups

Genotype distributions and allele frequencies for low and high BMI groups are presented in Tables 4.7 and 4.8. No deviations from Hardy-Weinberg expectations were evident in the data and no significant differences existed between the allele frequencies of the low and high BMI groups. There was a significant decline in the frequency of LDLR-P2 homozygotes from 64.1% in the low BMI group to 42.9% in the high BMI group (χ^2=6.97, d.f.=1, p=0.008). The corresponding decline in LDLR-P2 allele frequency from 0.802 ± 0.039 in low BMI to 0.690 ± 0.071 approached significance (Refer Table 4.8).

Table 4.7 Micronesian lipoprotein genotype frequencies in low and high body mass index groups.

Locus (Enzyme)	Genotype	Low			High		
		N	%	p†	N	%	p†
APOA1	M1-M1	8	11.8	0.422	5	20.8	0.880
(*Msp*I)	M1-M2	36	52.9		13	54.2	
	M2-M2	24	35.3		6	25.0	
	Total	68			24		
APOA2	M1-M1	35	46.7	0.968	15	57.7	0.318
(*Msp*I)	M1-M2	33	44.0		8	30.8	
	M2-M2	7	9.3		3	11.5	
	Total	75			26		
APOC1	D1-D1	52	76.5	0.107	13	59.1	1.000
(*Dra*I)	D1-D2	13	19.1		8	36.4	
	D2-D2	3	4.4		1	4.5	
	Total	68			22		
APOC2	T1-T1	18	23.7	0.602	7	28.0	0.208
(*Taq*I)	T1-T2	41	53.9		16	64.0	
	T2-T2	17	22.4		2	8.0	
	Total	76			25		
APOD	T1-T1	11	17.8	0.475	2	9.1	1.000
(*Taq*I)	T1-T2	26	41.9		9	40.9	
	T2-T2	25	40.3		11	50.0	
	Total	62			22		
APOE§	E*2/E*2	0	0.0	0.673	0	0.0	1.000
(*Cfo*I)	E*2/E*3	6	9.4		1	5.0	
	E*2/E*4	1	1.5		0	0.0	
	E*3/E*3	43	67.2		13	65.0	
	E*3/E*4	14	21.9		6	30.0	
	E*4/E*4	0	0.0		0	0.0	
	Total	64			20		
LDLR	P1-P1	2	3.8	1.000	1	4.7	0.611
(*Pvu*II)	P1-P2	17	32.1		11	52.4	
	P2-P2	34	64.1		9	42.9	
	Total	53			21		
LPL	P1-P1	35	66.0	1.000	14	66.7	1.000
(*Pvu*II)	P1-P2	16	30.2		6	28.6	
	P2-P2	2	3.8		1	4.7	
	Total	53			21		

† Testing Hardy-Weinberg equilibrium, χ^2 p value, d.f.=1, or exact probability.

§ APOE genotypes pooled into three classes for testing:- common homozygotes; common/rare heterozygotes; rare homozygotes and other heterozygotes.

Table 4.8 Micronesian lipoprotein allele frequencies in low and high body mass index groups.

Locus (Enzyme)	Allele	Low		High		$\chi^{2\dagger}$	p value
		Freq.	SE	Freq.	SE		
APOA1	M1	0.382	0.042	0.479	0.072	1.38	0.241
(*Msp*I)	M2	0.618	0.042	0.521	0.072		
	N	68		24			
APOA2	M1	0.687	0.038	0.731	0.061	0.36	0.551
(*Msp*I)	M2	0.313	0.038	0.269	0.061		
	N	75		26			
APOC1	D1	0.860	0.030	0.773	0.063	1.89	0.170
(*Dra*I)	D2	0.140	0.030	0.227	0.063		
	N	68		22			
APOC2	T1	0.507	0.041	0.600	0.069	1.32	0.251
(*Taq*I)	T2	0.493	0.041	0.400	0.069		
	N	76		25			
APOD	T1	0.387	0.044	0.295	0.069	1.18	0.278
(*Taq*I)	T2	0.613	0.044	0.705	0.069		
	N	62		22			
APOE	E*2	0.055	0.020	0.025	0.025	§	
(*Cfo*I)	E*3	0.828	0.033	0.825	0.060		
	E*4	0.117	0.028	0.150	0.056		
	N	64		20			
LDLR	P1	0.198	0.039	0.310	0.071	2.11	0.146
(*Pvu*II)	P2	0.802	0.039	0.690	0.071		
	N	53		21			
LPL	P1	0.811	0.038	0.810	0.061	0.00	1.000
(*Pvu*II)	P2	0.189	0.038	0.190	0.061		
	N	53		21			

† Contingency χ^2 of allele counts, d.f.=1, excluding APOE, d.f.=2.

§ Pooled allele count tests for APOE:- E*2 vs E*3 and E*4, p=0.682; E*3 vs E*2 and E*4, p=0.846; E*4 vs E*2 and E*3, p=0.784.

4.4.1.4 Allele frequencies and genotype distributions in low and high waist-to-hip ratio groups

Three of the genotype distributions examined in relation to high or low waist-to-hip ratio differed significantly from Hardy-Weinberg expectations (Table 4.9). APOD heterozygotes (32.8%) were under-represented in comparison to an expectation of 46.1% in the low WHR group. The reverse was true for the high WHR group where the frequency of APOD heterozygotes (68.2%) exceeded that expected (44.9%) by some 23.3%. There was an excess of 13.0% over expectation in the observed frequency of heterozygotes for APOC2 in the low WHR group with a concomitant deficit in both homozygote classes.

An increase in the frequency of the APOC2-T1 allele from 0.493 ± 0.041 in the low WHR group to 0.635 ± 0.073 in the high WHR group was evident (Table 4.10). Although this increase itself bordered on significance, the changes in APOC2-T1 homozygote frequencies from 17.8% in the low group to 46.2% in the high group and in APOC2 heterozygote frequencies from 63.0% in the low group to 34.6% in the high group were both significant (χ^2=6.73, d.f.=1, p=0.009 and χ^2=5.16, d.f.=1, p=0.023 respectively). The frequency of the APOE*2 allele was 0.032 ± 0.016 in the low WHR group and 0.105 ± 0.050 in the high WHR group. This increase approached significance at p=0.085.

Table 4.9 Micronesian lipoprotein genotype frequencies in low and high waist-to-hip ratio groups.

Locus (Enzyme)	Genotype	Low			High		
		N	%	p†	N	%	p†
APOA1 (*Msp*I)	M1-M1	10	14.7	0.705	3	13.6	0.682
	M1-M2	35	51.5		12	54.6	
	M2-M2	23	33.8		7	31.8	
	Total	68			22		
APOA2 (*Msp*I)	M1-M1	36	48.7	0.916	13	52.0	0.631
	M1-M2	32	43.2		9	36.0	
	M2-M2	6	8.1		3	12.0	
	Total	74			25		
APOC1 (*Dra*I)	D1-D1	47	70.1	0.387	17	77.3	0.324
	D1-D2	17	25.4		4	18.2	
	D2-D2	3	4.5		1	4.5	
	Total	67			22		
APOC2 (*Taq*I)	T1-T1	13	17.8	0.040	12	46.2	0.330
	T1-T2	46	63.0		9	34.6	
	T2-T2	14	19.2		5	19.2	
	Total	73			26		
APOD (*Taq*I)	T1-T1	12	19.7	0.041	0	0.0	0.048
	T1-T2	20	32.8		15	68.2	
	T2-T2	29	47.5		7	31.8	
	Total	61			22		
APOE§ (*Cfo*I)	E*2/E*2	0	0.0	0.189	0	0.0	1.000
	E*2/E*3	4	6.3		3	15.8	
	E*2/E*4	0	0.0		1	5.2	
	E*3/E*3	42	66.7		12	63.2	
	E*3/E*4	17	27.0		3	15.8	
	E*4/E*4	0	0.0		0	0.0	
	Total	63			19		
LDLR (*Pvu*II)	P1-P1	3	5.6	0.712	0	0.0	0.277
	P1-P2	18	34.0		10	50.0	
	P2-P2	32	60.4		10	50.0	
	Total	53			20		
LPL (*Pvu*II)	P1-P1	37	69.8	1.000	11	55.0	0.591
	P1-P2	15	28.3		7	35.0	
	P2-P2	1	1.9		2	10.0	
	Total	53			20		

† Testing Hardy-Weinberg equilibrium, χ^2 p value, d.f.=1, or exact probability.
§ APOE genotypes pooled into three classes for testing:- common homozygotes; common/rare heterozygotes; rare homozygotes and other heterozygotes.

Table 4.10 Micronesian lipoprotein allele frequencies in low and high waist-to-hip ratio groups.

Locus (Enzyme)	Allele	Low		High		$\chi^{2\dagger}$	p value
		Freq.	SE	Freq.	SE		
APOA1	M1	0.404	0.042	0.409	0.074	0.00	0.956
(*Msp*I)	M2	0.596	0.042	0.591	0.074		
	N	68		22			
APOA2	M1	0.703	0.038	0.700	0.065	0.00	0.971
(*Msp*I)	M2	0.297	0.038	0.300	0.065		
	N	74		25			
APOC1	D1	0.828	0.033	0.864	0.052	0.30	0.583
(*Dra*I)	D2	0.172	0.033	0.136	0.052		
	N	67		22			
APOC2	T1	0.493	0.041	0.635	0.073	3.08	0.079
(*Taq*I)	T2	0.507	0.041	0.365	0.073		
	N	73		26			
APOD	T1	0.361	0.043	0.341	0.071	0.06	0.815
(*Taq*I)	T2	0.639	0.043	0.659	0.071		
	N	61		22			
APOE	E*2	0.032	0.016	0.105	0.050	§	
(*Cfo*I)	E*3	0.833	0.033	0.789	0.066		
	E*4	0.135	0.030	0.105	0.050		
	N	63		19			
LDLR	P1	0.226	0.041	0.250	0.068	0.09	0.764
(*Pvu*II)	P2	0.774	0.041	0.750	0.068		
	N	53		20			
LPL	P1	0.840	0.036	0.725	0.071	2.46	0.117
(*Pvu*II)	P2	0.160	0.036	0.275	0.071		
	N	53		20			

† Contingency χ^2 of allele counts, d.f.=1, excluding APOE, d.f.=2.

§ Pooled allele count tests for APOE:- E*2 vs E*3 and E*4, p=0.085; E*3 vs E*2 and E*4, p=0.705; E*4 vs E*2 and E*3, p=0.785.

4.4.1.5 Phenotypic and allelic means

The phenotypic means for CHO, TG, BMI and WHR have been calculated for all eight polymorphic sites and are presented in Tables 4.11, 4.12 and 4.13. Also displayed in these tables are the p values for one way analysis of variance between phenotypic means, and p values for *t*-tests between the means for the presence or absence of an allele. In the case of APOE the allelic means are also presented in tabulated form (Table 4.14).

At the APOA1 locus the presence of the APOA1-M1 was significantly associated with an increase in BMI of 3.35 ± 1.70 kg/m^2. This increase equated to 9.7% of the average BMI for the sample. No significant correlations of BMI with other variables were evident after multiple regression analysis.

The presence of the APOC2-T2 allele was significantly associated with decreased plasma triglyceride levels. The mean plasma triglyceride concentration was 1.26 ± 0.09 mM in the presence of the T2 allele and increased to 1.85 ± 0.29 mM in its absence (p=0.058). This increase of 0.59 ± 0.22 was equivalent to 43.1% of the average plasma triglyceride concentration in the sample. The differences across phenotypic means for this RFLP were also significant (p=0.032 and refer Table 4.11).

Plasma triglyceride levels were strongly correlated with plasma cholesterol levels and correlated to a lesser extent with diabetic status and WHR in the sample population (Section 4.3.6). To examine the strength of the plasma

triglyceride to APOC2 association, transformed triglyceride values were regressed on cholesterol levels, age, sex, BMI, WHR, and diabetic status (Table 4.15). The plasma cholesterol levels and WHR had a significant effect on transformed plasma triglyceride levels (CHO β=0.514, ChR^2=0.290; WHR β=0.262, ChR^2=0.068). When plasma triglyceride levels were adjusted for cholesterol level and WHR the association between TG levels and APOC2 was no longer significant. When plasma cholesterol levels were adjusted for the effect of plasma triglyceride levels (Table 4.15, β=0.538, ChR^2=0.290) a significant association between the presence of the APOC2-T2 allele and decreased plasma cholesterol concentrations was observed (p=0.035). The effect of the presence of APOC2-T2 was to lower the adjusted cholesterol level 6.2% from 5.47 ± 0.14 mM to 5.13 ± 0.07 mM. This association results from the strong correlation seen between triglyceride levels and cholesterol levels within the sample; there was no primary relationship between APOC2 and cholesterol as allele frequencies for APOC2 did not differ between low and high cholesterol groups.

The APOC2 polymorphism also had an affect on waist-to-hip ratios. A marginal association of APOC2 phenotype to mean WHR was evident (p=0.075 and refer Table 4.11). The majority of this difference was due to the fall in WHR in the presence of the APOC2-T2 allele. The allelic mean in the presence of APOC2-T2 was 0.86 ± 0.01 and 0.89 ± 0.01 in its absence (p=0.025). This increase represented 3.5% of the mean WHR measurement for the total sample.

Waist-to-hip ratios were also increased in the presence of APOD-T2 (p=0.033). This increase was marginally significant across phenotypes (p=0.074, Table 4.12) and represented 4.7% of the mean WHR for the sample as a whole.

As WHR was correlated with age, diabetic status, sex and plasma triglycerides in the total sample (Section 4.3.6) a multiple regression was performed in order to adjust the WHR measurements for the sample for these covariates (Table 4.15). Other variables, including plasma CHO levels and BMI were presented for entry, but only those already known to be correlated with WHR entered the regression equation (sex, diabetes, tranTG and age). After adjustment for these variables the associations between WHR and APOC2, and WHR and APOD were no longer significant.

The presence of the APOE*2 and APOE*4 alleles did not significantly affect the mean values for CHO, TG, BMI or WHR (Table 4.14).

Table 4.11 Plasma cholesterol, plasma triglycerides, body mass index and waist-to-hip ratio phenotypic means for APOA1, APOA2, APOC1 and APOC2.

Phenotype	1-1			1-2			2-2			p values†		
	$\bar{x}$#	SE	N	$\bar{x}$	SE	N	$\bar{x}$	SE	N	Phen.§	Pres. 1	Pres. 2
APOA1 (*Msp*I)												
CHO	5.03	0.29	13	5.17	0.17	49	5.36	0.18	30	0.676	0.433	0.536
TG	1.09	0.19	13	1.41	0.13	49	1.58	0.24	30	0.355	0.364	0.216
BMI	37.93	2.13	13	35.57	1.18	49	32.71	1.18	30	0.093	0.051	0.138
WHR	0.87	0.02	13	0.86	0.01	47	0.87	0.01	30	0.779	0.744	0.612
APOA2 (*Msp*I)												
CHO	5.24	0.14	50	5.29	0.19	41	4.82	0.23	9	0.497	0.242	0.855
TG	1.39	0.14	50	1.44	0.17	41	1.35	0.30	9	0.964	0.863	0.887
BMI	35.49	1.09	50	34.23	1.08	41	36.02	2.61	10	0.656	0.659	0.538
WHR	0.87	0.01	49	0.85	0.01	41	0.86	0.02	9	0.217	0.896	0.089
APOC1 (*Dra*I)												
CHO	5.30	0.14	65	5.61	0.22	21	5.15	....	4	.....	.....	0.356
TG	1.50	0.14	65	1.36	0.19	21	1.94	....	4	.....	.....	0.838
BMI	34.26	0.89	65	36.21	1.96	21	35.80	....	4	.....	.....	0.298
WHR	0.86	0.01	64	0.87	0.01	21	0.85	....	4	.....	.....	0.662
APOC2 (*Taq*I)												
CHO	5.25	0.26	25	5.25	0.12	56	4.99	0.29	19	0.639	0.342	0.807
TG	1.85	0.29	25	1.28	0.10	56	1.19	0.17	19	0.032	0.317	0.058
BMI	35.54	1.59	25	35.23	1.07	57	32.54	1.01	19	0.350	0.054	0.574
WHR	0.89	0.01	25	0.86	0.01	55	0.85	0.02	19	0.075	0.305	0.025

† Significance tests not valid where N<5.
§ Abbreviations for p values:- Phen. = ANOVA between phenotypic means; Pres. 1 = *t*-test between means for presence of allele 1 against absence of allele 1; Pres. 2 = *t*-test between means for presence of allele 2 against absence of allele 2.
# CHO (mM), TG (mM), BMI (kg/m^2).

Table 4.12 Plasma cholesterol, plasma triglycerides, body mass index and waist-to-hip ratio phenotypic means for APOD, LDLR and LPL.

Phenotype	1-1			1-2			2-2			p values†		
	$\bar{x}$#	SE	N	$\bar{x}$	SE	N	$\bar{x}$	SE	N	Phen.§	Pres. 1	Pres. 2
APOD (*Taq*I)												
CHO	4.89	0.17	13	5.14	0.16	35	5.24	0.22	36	0.619	0.520	0.190
TG	1.44	0.27	13	1.48	0.16	35	1.38	0.20	36	0.918	0.696	0.988
BMI	33.75	1.42	13	34.44	1.38	35	36.12	1.44	36	0.553	0.290	0.523
WHR	0.83	0.02	12	0.88	0.01	35	0.86	0.01	36	0.074	0.991	0.033
LDLR (*Pvu*II)												
CHO	4.79	....	3	5.34	0.24	28	5.11	0.19	43	.....	0.544	.....
TG	0.71	....	3	1.52	0.24	28	1.45	0.16	43	.....	0.876	.....
BMI	33.47	....	3	36.87	1.68	28	33.84	1.09	43	.....	0.149	.....
WHR	0.79	....	3	0.88	0.01	28	0.85	0.01	42	.....	0.220	.....
LPL (*Pvu*II)												
CHO	5.28	0.17	49	4.91	0.26	22	5.66	....	3	.....	.....	0.349
TG	1.59	0.18	49	1.25	0.17	22	1.01	....	3	.....	.....	0.118
BMI	34.62	1.21	49	35.83	1.46	22	34.40	....	3	.....	.....	0.597
WHR	0.86	0.01	48	0.86	0.02	22	0.91	....	3	.....	.....	0.514

† Significance tests not valid where N<5.
§ Abbreviations for p values:- Phen. = ANOVA between phenotypic means; Pres. 1 = *t*-test between means for presence of allele 1 against absence of allele 1; Pres. 2 = *t*-test between means for presence of allele 2 against absence of allele 2.
# CHO (mM), TG (mM), BMI (kg/m^2).

Table 4.13 Plasma cholesterol, plasma triglycerides, body mass index and waist-to-hip ratio phenotypic means for APOE.

Phenotype†	APOE*2/APOE*3			APOE*2/APOE*4			APOE*3/APOE*3			APOE*3/APOE*4		
	$\bar{x}$§	SE	N	$\bar{x}$	SE	N	$\bar{x}$	SE	N	$\bar{x}$	SE	N
CHO	5.01	0.81	7	5.31	....	1	5.15	0.12	56	5.31	0.22	20
TG	1.81	0.80	7	3.25	....	1	1.35	0.12	56	1.24	0.16	20
BMI	33.63	2.29	7	27.80	....	1	34.81	0.95	56	35.96	1.81	20
WHR	0.89	0.22	7	0.91	....	1	0.86	0.01	56	0.86	0.02	20

† Phenotypes APOE*2/APOE*2 and APOE*4/APOE*4 not represented in the sample.
§ CHO (mM), TG (mM), BMI (kg/m^2).

Table 4.14 Plasma cholesterol, plasma triglycerides, body mass index and waist-to-hip ratio allelic means for APOE.

	Allele present			Allele absent			
	$\bar{x}$[†]	SE	N	x	SE	N	p value[§]
APOE*2							
CHO	5.05	0.70	8	5.19	0.11	76	0.849
TG	1.99	0.72	8	1.32	0.20	76	0.385
BMI	32.90	2.11	8	35.12	0.84	76	0.412
WHR	0.89	0.02	8	0.86	0.01	74	0.144
APOE*3							
CHO	5.17	0.12	83	5.31	....	1	.....
TG	1.36	0.11	83	3.25	....	1	.....
BMI	35.00	0.79	73	27.80	....	1	.....
WHR	0.86	0.01	81	0.91	....	1	.....
APOE*4							
CHO	5.31	0.21	21	5.13	0.14	63	0.491
TG	1.34	0.18	21	1.40	0.14	63	0.805
BMI	35.60	1.77	21	34.68	0.88	63	0.615
WHR	0.86	0.01	21	0.86	0.01	61	0.919

† CHO (mM), TG (mM), BMI (kg/m^2).
§ Two-tailed *t*-test, not valid where N<5.

Table 4.15 Multiple regression of plasma triglyceride, plasma cholesterol levels, and waist-to-hip ratios.

Independent variable	β	SE	FCh[†]	SigFCh	ChR^{2}[§]
Dependent variable transformed plasma triglyceride level					
N=113					
CHO	0.514	0.077	45.23	0.000	0.290
WHR	0.262	0.077	11.64	0.001	0.068
Dependent variable plasma cholesterol level					
N=113					
tranTG	0.538	0.080	45.23	0.000	0.290
Dependent variable waist-to-hip ratio					
N=113					
Sex	0.415	0.077	31.00	0.000	0.218
Diabetes	0.190	0.083	14.19	0.000	0.089
tranTG	0.199	0.078	5.86	0.017	0.035
Age	0.201	0.083	5.89	0.017	0.034

† Change in F upon addition of new independent variable.
§ Proportion of variance explained by each independent variable.

4.4.2 Micronesian lipoprotein genes in non-insulin dependent diabetes mellitus - population study

4.4.2.1 Genotype distributions and allele frequencies

Genotype distributions and allele frequencies for the eight loci under study are given in Tables 4.16 and 4.17. All of the genotype distributions examined in relation to NIDDM were consistent with expectations under Hardy-Weinberg equilibrium except APOC2. The expected frequency of APOC2 heterozygotes in the patient group was 49.1% and the observed frequency was some 13.6% higher at 62.7%. This increase mirrored the increase seen in observed over expected heterozygote frequencies in the low WHR group and reported in Section 4.4.1.4.

APOE allele frequencies differed marginally between patient and control groups. The rare alleles APOE*2 and APOE*4 were present in 17 of 36 patients (47.2%) and only 15 of 54 controls (27.8%) (χ^2=2.77, d.f.=1, p=0.096). The larger portion of this difference was due to an increase in the frequency of the APOE*2 allele from 0.028 ± 0.016 in controls to 0.097 ± 0.035 in patients. There was an attendant marginally significant decline in APOE*3 frequencies from 0.861 ± 0.033 in the control group to 0.750 ± 0.051 in the patients (χ^2=2.86, d.f.=1, p=0.091).

The presence of the APOD-T2 allele was significantly associated with NIDDM. The frequency of this allele increased from 0.594 ± 0.048 in the control group to 0.777 ± 0.039 in the patient group (χ^2=8.45, d.f.=1, p=0.004). The corresponding decline in the APOD-T1 homozygote

frequency from 20.8% in controls to 3.6% in patients was highly significant (Fisher's exact test p=0.007). There was also marginal significance in the increase in APOD-T2 homozygosity from 39.6% in controls to 58.9% in patients (χ^2=3.32, d.f.=1, p=0.068).

Table 4.16 Micronesian lipoprotein genotype frequencies in diabetes patients and control populations.

Locus (Enzyme)	Genotype	Patients			Controls		
		N	%	p†	N	%	p†
APOA1	M1-M1	11	20.0	1.000	8	13.6	0.591
(*Msp*I)	M1-M2	27	49.1		31	52.5	
	M2-M2	17	30.9		20	33.9	
	Total	55			59		
APOA2	M1-M1	27	47.4	0.960	30	46.9	0.679
(*Msp*I)	M1-M2	25	43.8		26	40.6	
	M2-M2	5	8.8		8	12.5	
	Total	57			64		
APOC1	D1-D1	40	72.7	0.596	43	75.4	0.285
(*Dra*I)	D1-D2	13	23.7		12	21.1	
	D2-D2	2	3.6		2	3.5	
	Total	55			57		
APOC2	T1-T1	15	25.4	0.053	16	25.4	0.999
(*Taq*I)	T1-T2	37	62.7		32	50.8	
	T2-T2	7	11.9		15	23.8	
	Total	59			63		
APOD	T1-T1	2	3.6	0.714	11	20.8	0.278
(*Taq*I)	T1-T2	21	37.5		21	39.6	
	T2-T2	33	58.9		21	39.6	
	Total	56			53		
APOE§	E*2/E*2	0	0.0	0.403	0	0.0	0.575
(*Cfo*I)	E*2/E*3	6	16.6		3	5.6	
	E*2/E*4	1	2.8		0	0.0	
	E*3/E*3	19	52.8		39	72.2	
	E*3/E*4	10	27.8		12	22.2	
	E*4/E*4	0	0.0		0	0.0	
	Total	36			54		
LDLR	P1-P1	2	5.3	0.691	2	4.2	1.000
(*Pvu*II)	P1-P2	17	44.7		15	31.2	
	P2-P2	19	50.0		31	64.6	
	Total	38			48		
LPL	P1-P1	30	63.8	0.200	30	62.5	1.000
(*Pvu*II)	P1-P2	13	27.7		16	33.3	
	P2-P2	4	8.5		2	4.2	
	Total	47			48		

† Testing Hardy-Weinberg equilibrium, χ^2 p value, d.f.=1, or exact probability.
§ APOE genotypes pooled into three classes for testing:- common homozygotes; common/rare heterozygotes; rare homozygotes and other heterozygotes.

Table 4.17 Micronesian lipoprotein allele frequencies in diabetes patients and control populations.

Locus (Enzyme)	Allele	Patients		Controls		$\chi^{2\dagger}$	p value
		Freq.	SE	Freq.	SE		
APOA1	M1	0.445	0.047	0.398	0.045	0.52	0.471
(*Msp*I)	M2	0.555	0.047	0.602	0.045		
	N	55		59			
APOA2	M1	0.693	0.043	0.672	0.041	0.12	0.725
(*Msp*I)	M2	0.307	0.043	0.328	0.041		
	N	57		64			
APOC1	D1	0.845	0.035	0.860	0.032	0.09	0.764
(*Dra*I)	D2	0.155	0.035	0.140	0.032		
	N	55		57			
APOC2	T1	0.568	0.046	0.508	0.045	0.88	0.349
(*Taq*I)	T2	0.432	0.046	0.492	0.045		
	N	59		63			
APOD	T1	0.223	0.039	0.406	0.048	8.45	0.004
(*Taq*I)	T2	0.777	0.039	0.594	0.048		
	N	56		53			
APOE	E*2	0.097	0.035	0.028	0.016	§	
(*Cfo*I)	E*3	0.750	0.051	0.861	0.033		
	E*4	0.153	0.042	0.111	0.030		
	N	36		54			
LDLR	P1	0.276	0.051	0.198	0.041	1.46	0.227
(*Pvu*II)	P2	0.724	0.051	0.802	0.041		
	N	38		48			
LPL	P1	0.777	0.043	0.792	0.041	0.06	0.801
(*Pvu*II)	P2	0.223	0.043	0.208	0.041		
	N	47		48			

† Contingency χ^2 of allele counts, d.f.=1, excluding APOE, d.f.=2.

§ Pooled allele count tests for APOE:- E*2 vs E*3 and E*4, p=0.092; E*3 vs E*2 and E*4, p=0.091; E*4 vs E*2 and E*3, p=0.554.

4.4.3 Multivariate analysis of Micronesian lipoprotein genes in non-insulin dependent diabetes mellitus, lipoproteinemia and obesity

4.4.3.1 Lipoprotein genes in obesity and lipoproteinemia - multiple regression

Four sets of multiple regression analyses were undertaken with transformed plasma triglyceride concentrations, plasma cholesterol, body mass index and waist-to-hip ratios as successive dependent variables. The analyses were performed on the random series data set with the sample size limited to the number of individuals genotyped at each of the eight loci examined. In each analysis the variables listed below (excluding the dependent variable under examination) were presented for entry to the regression as independent variables:-

age (years)
sex
diabetic status
transformed plasma triglyceride level (mM)
plasma cholesterol level (mM)
body mass index (kg/m^2)
waist-to-hip ratio

There are known correlations between some of these variables (CHO and TG, WHR and sex for example, refer Section 4.3.6). It was therefore expected that collinearity would be present between some of the independent variables. However all independent variables were subject to a tolerance test before entry to the regression, and none were rejected on the basis of collinearity alone.

In addition to the variables listed above, dummy variables for the presence of each allele of the eight lipoprotein genes under study were presented in successive runs of the analyses. The individual presentation of the loci was

necessary to ensure that the same cases were used to estimate all coefficients within one analysis, and because significance levels obtained from pairwise matrices may be unreliable (SPSS Inc., 1986). An interactive term for age and diabetic status was examined, but the significance of this interaction did not reach the inclusion threshold of 5% in any of the preliminary analyses and so was not considered further.

In the analysis of CHO and BMI none of the loci tested were found to have a significant effect on the dependent variable. The presence of APOC2-T2 had a significant effect on transformed plasma triglyceride levels (β=-0.177, ChR^2=0.029, Table 4.18). This effect was less than that of either plasma cholesterol concentrations, or WHR. The proportion of triglyceride variance explained by the model was 40.2%, with the proportion explained by CHO levels, WHR and the presence of APOC2-T2 being 30.3%, 7.0% and 2.9% respectively.

The APOC2-T2 allele also had a marginally significant effect on WHR (Table 4.18). With the sample size limited by the number of individuals genotyped for APOC2 (N=95) diabetic status was no longer a significant independent variable for WHR (β=0.132, ChR^2=0.085, p=0.149 for entry, compare with Section 4.4.1.5, Table 4.15). The effect of the APOC2-T2 allele on WHR was minor (β=-0.147, ChR^2=0.020) in comparison to the effects of age, sex and plasma triglyceride level (Table 4.18).

Table 4.18 Multiple regression of plasma triglyceride, and waist-to-hip ratios - effect of APOC2-T2.

Independent variable	β	SE	FCh†	SigFCh	ChR²§
Dependent variable transformed plasma triglyceride level					
N=95					
CHO	0.492	0.083	40.40	0.000	0.303
WHR	0.226	0.086	10.29	0.002	0.070
APOC2-T2	-0.177	0.084	4.45	0.038	0.029
Dependent variable waist-to-hip ratio					
N=95					
Age	0.350	0.082	21.03	0.000	0.184
Sex	0.329	0.084	19.07	0.000	0.140
tranTG	0.226	0.086	9.95	0.002	0.067
APOC2-T2	-0.147	0.085	3.04	0.085	0.020

† Change in F upon addition of new independent variable.
§ Proportion of variance explained by each independent variable.

4.4.3.2 Lipoprotein genes in non-insulin dependent diabetes mellitus - loglinear analysis

The independent variables of age, sex, transformed plasma triglyceride concentration, plasma cholesterol concentration, BMI and WHR were tested for successive significant entry to a loglinear model with diabetic status as a binary dependent variable. Interactive terms between the variables listed above were also tested for entry. These analyses were performed on the NIDDM series data set. The total sample size in each analysis was limited to the number of individuals genotyped for the locus under investigation.

An initial deviance of 187.48 (d.f.=136) was reduced to 148.51 (d.f.=133) with the successive addition of WHR (% Var=12.03), age (% Var=5.68) and a WHR*age interaction term (% Var=3.08) to the model (Table 4.19). This model accounted for 20.79% of the variability in diabetic status.

Transformed triglyceride plasma levels, while being a significant predictor of diabetic status when presented alone, did not enter the model after the predictors of WHR, age and WHR*age had been entered.

Table 4.19 Loglinear model of diabetic status.

Independent variable	Model deviance	Change in deviance	d.f.	p†	% Var§
Dependent variable diabetic status					
Empty model	187.48	.....	136	.....	....
Successive entry					
WHR	164.92	22.56	1	<0.001	12.0
Age	154.28	10.64	1	0.001	5.7
WHR*age	148.51	5.77	1	0.016	3.1
Combined entry					
WHR, age and WHR*age	148.51	38.97	3	<0.001	20.8

† p value for change in deviance, distributed as χ^2 with associated degrees of freedom.
§ % of variance = change in deviance over empty model deviance.

Genotypes, or the presence of alleles for the eight loci were first presented individually for entry to the model; if the genotype or allele presence was found to be a significant predictor for diabetic status they were presented to a model which already included WHR, age and WHR*age and the results were re-examined.

The genotypes and the presence of either allele for APOD was found to be a significant predictor of diabetic status (Table 4.20; APOD genotype % Var.=5.5; APOD-T1 % Var.=2.6; APO-T2 % Var=4.7). Once the more powerful predictors of diabetic status were included in the model the effect of the APOD locus could no longer be detected (Table 4.20).

Table 4.20 Loglinear model of diabetic status - effect of APOD genotypes and alleles.

Independent variable	Model deviance	Change in deviance	d.f.	p†	% Var§
Dependent variable diabetic status					
Empty model	146.91	....	105	.....	...
Lone entry					
APOD genotype	138.80	8.11	2	0.017	5.5
APOD-T1	143.11	3.80	1	0.051	2.6
APOD-T2	140.06	6.85	1	0.009	4.7
Combined entry					
WHR, age and WHR*age	117.64	29.27	3	<0.001	19.9
*Lone entry after WHR, age and WHR*age*					
APOD genotype	114.62	3.02	2	0.221	2.1
APOD-T1	115.12	2.52	1	0.112	1.7
APOD-T2	116.25	1.39	1	0.238	1.0

† p value for change in deviance, distributed as χ^2 with associated degrees of freedom.
§ % of variance = change in deviance over empty model deviance.

APOC2-T1 was found to have a marginal effect on diabetic status (Table 4.21; % Var=1.8), the significance of which was lowered further when the model included WHR, age and WHR*age. The presence of APOE*2 allele was a significant predictor of diabetic status (Table 4.21; % Var=3.8), but not after the stronger predictors of WHR, age and WHR*age had been entered into the model.

Table 4.21 Loglinear model of diabetic status - effect of APOC2-T1 and APOE*2.

Independent variable	Model deviance	Change in deviance	d.f.	p†	% Var§
Dependent variable diabetic status					
Empty model	163.45	....	117	.....	...
Lone entry					
APOC2-T1	160.44	3.01	1	0.083	1.8
Combined entry					
WHR, age and WHR*age	129.86	33.59	3	<0.001	20.6
*Lone entry after WHR, age and WHR*age*					
APOC2-T1	127.80	2.06	1	0.150	1.3
Dependent variable diabetic status					
Empty model	116.42	....	86	.....	...
Lone entry					
APOE*2	111.99	4.43	1	0.035	3.8
Combined entry					
WHR, age and WHR*age	85.55	30.88	3	<0.001	26.5
*Lone entry after WHR, age and WHR*age*					
APOE*2	82.93	2.62	1	0.106	2.3

† p value for change in deviance, distributed as χ^2 with associated degrees of freedom.
§ % of variance = change in deviance over empty model deviance.

4.5 DISCUSSION

4.5.1 APOC2 in non-insulin dependent diabetes mellitus and hypertriglyceridemia

A significant association was found between plasma triglyceride levels and the *Taq*I polymorphic site of APOC2. A dosage effect was apparent with the presence of the APOC2-T2 allele (3.8kb) being associated with a 31.9% decrease in plasma triglyceride concentrations. Once transformed triglyceride levels were adjusted for plasma cholesterol concentrations and WHR this dosage effect was no longer significant. However, it is more appropriate to

evaluate the confounding effects of cholesterol and waist-to-hip ratios on the APOC2 triglyceride association by means of multivariate analysis. These analyses showed that the APOC2-T2 allele had an influence on triglyceride variability after adjustment for the influences of cholesterol levels and waist-to-hip ratios. The APOC2-triglyceride relationship was also consistent with a tendency, although non-significant, towards an excess of APOC2-T1 homozygotes (3.5kb) in individuals in the upper quartile for plasma triglyceride levels.

The relationship between the APOC2 *Taq*I RFLP and plasma triglyceride levels were replicated, in the main, with respect to waist-to-hip ratios. There was a marginally significant dosage effect of APOC2-T2 in lowering waist-to-hip ratios, and a significant decrease in the waist-to-hip ratios of individuals who possessed the APOC2-T2 allele. There was an attendant increase of APOC2-T1 homozygotes in individuals with waist-to-hip ratios below the top quartile for the population. In multiple regression analyses the APOC2-T2 allele was seen to influence waist-to-hip ratios, although the significance of this influence was marginal.

The increase of plasma triglyceride levels in NIDDM is well documented (Betteridge, 1986; Howard, 1987) and was evident in the present study (Section 4.3.6). However, the significance of the effect of the APOC2 locus as a predictor for NIDDM was marginal. It would appear then, that the primary association of the APOC2 locus in the

Micronesian population is with alterations in plasma triglyceride concentrations and not NIDDM or WHR.

The apolipoprotein C2 protein is a cofactor for lipoprotein lipase, the enzyme responsible for hydrolysis of plasma triglycerides to free fatty acids and monoacylglycerols (Herbert *et al.*, 1983). Individuals with an apolipoprotein C2 deficiency often present clinically with a massive accumulation of chylomicrons and triglycerides, the classical symptoms of lipoprotein lipase deficiency (type I hyperlipoproteinemia) (Yamamura *et al.*, 1979). It is possible that alterations in the efficiency of action of apolipoprotein C2, rather than its complete absence, could also impair LPL function and lead to increases in plasma triglycerides.

The APOC2-T1 allele is reported by Deeb *et al.* (1986) to be associated with increased cholesterol levels. Different populations with different linkage disequilibrium relationships may show different RFLP associations with triglyceride concentrations, or its correlate, cholesterol levels. However, the marked effect of allelic variation at the APOC2 locus on triglyceride levels observed in this study has not been reported previously.

4.5.2 APOD in non-insulin dependent diabetes mellitus and hypercholesterolemia

The presence of the APOD-T2 allele (*Taq*I 2.7kb) was significantly associated with NIDDM. The allele was a significant lone predictor of diabetic status as indicated by loglinear modelling. The strength of the relationship

between APOD and NIDDM was minor in comparison to the relationships of the stronger correlates of age and WHR with NIDDM. Average waist-to-hip ratios were increased in the presence of the APOD-T2 allele, but possibly due only to the strong association between diabetic status and WHR. The observation of the association between APOD alleles and NIDDM is unique, and the possible contribution of the APOD polymorphism to NIDDM susceptibility has not been examined in other populations (Refer Section 4.2.2).

As a result of the increase in plasma cholesterol concentrations in patients with NIDDM the APOD-T2 allele was also found at increased frequency in individuals with plasma cholesterol levels in the top quartile of the population. Although the presence of the APOD-T2 allele was associated with increased cholesterol levels in the general population, these increases were not substantial enough to gain significance.

Lecithin-cholesterol acyltransferase catalyzes the esterification of free cholesterol in the plasma, forming cholesteryl esters. Apolipoproteins A1, C1, A4 and E are known to activate LCAT to different degrees, while apolipoproteins A2, C2 and C3 have an inhibitory effect on this enzyme (Refer Table 1.2). Steyrer and Kostner, (1988) postulate that apolipoprotein D may have a stabilizing effect on LCAT. However, the degree to which this suggested stabilization may result in altered plasma cholesterol levels is unknown.

The associations apparent between the APOD locus and NIDDM may be the result of linkage disequilibrium with another gene. A possible candidate is the glucose transporter gene (GLUT2, liver isoform) which has been localized to 3q14.2-q26.3 (Fukumoto *et al.*, 1989). This locality lies within the range 3p14.2-qter, so far determined for the position of the APOD locus (Drayna *et al.*, 1987*a*). The GLUT2 gene has not been examined in the Micronesian population.

4.5.3 APOE in non-insulin dependent diabetes mellitus and hyperlipoproteinemia

In contradiction with generally accepted patterns of associations between APOE alleles and hyperlipoproteinemia (Assmann *et al.*, 1984; Sing and Davignon, 1985; Ehnholm *et al.*, 1986), no associations were seen in this population between the APOE alleles and either hypertriglyceridemia or hypercholesterolemia.

The frequency of the APOE*2 allele was marginally increased in NIDDM and in individuals with high waist-to-hip ratios. This observation was in agreement with a previous report of increased APOE*2 containing phenotypes in NIDDM (Vogelberg and Maucy, 1988). Restricted sample sizes have precluded the division of the sample into patients or controls with and without obesity, or with or without increased plasma lipoproteins levels. However, given that the average BMI was high among both patients with diabetes and controls (35.17 ± 1.04 and 34.80 ± 0.77 respectively), and that only two of the diabetics and two of the controls had BMI values of less than 24.00, it is obvious that comparisons between

non-obese and obese individuals *per se* in this population, would be impossible. The sample then is of obese individuals with and without NIDDM. There was no evidence to support the suggestion of Fumeron *et al.* (1988) that the APOE*4 allele is more likely to be associated with hypertriglyceridemia in obese individuals.

4.6 CONCLUSIONS

1) The presence of the APOC2 *Taq*I 3.8kb allele was significantly associated with lowered plasma triglyceride levels in the Micronesian population.

2) The APOC2 *Taq*I 3.8kb allele was also associated with lower waist-to-hip ratios in the Micronesian population. This association, may however be secondary to that of the effect of APOC2 on triglyceride levels.

3) The APOD *Taq*I 2.7kb allele was significantly associated with non-insulin dependent diabetes mellitus in the Micronesian population and was more common in individuals with high cholesterol levels. The primary association, however, appears to be with non-insulin dependent diabetes mellitus and not plasma cholesterol levels.

4) The APOD *Taq*I 2.7kb allele was associated with increased waist-to-hip ratios. This association is believed to be secondary to the associations of the APOD RFLP with non-insulin dependent diabetes mellitus.

5) There is limited evidence for an association between the APOE*2 allele and non-insulin dependent diabetes mellitus. There is, however, no support for the often reported associations between APOE*2 and hypertriglyceridemia or APOE*4 and hypercholesterolemia.

SECTION 5

LINKAGE OF MICRONESIAN LIPOPROTEIN GENES WITH LIPOPROTEINEMIA AND DIABETIC STATUS

5.1 AIMS

1) To test for linkage between seven lipoprotein restriction fragment length polymorphisms, the common apolipoprotein E alleles and
   a) liability for non-insulin dependent diabetes mellitus
   b) hypercholesterolemia
   c) hypertriglyceridemia.

5.2 INTRODUCTION

5.2.1 Multifactorial disorders and linkage analysis

Inherited diseases with clear modes of inheritance, known biochemical defects, early age of onset, and full penetrance are most amenable to genetic analysis. However, the genes for some age dependent single gene disorders such as polyposis coli, polycystic kidney disease and multiple endocrine neoplasia have been mapped successfully (Reeders *et al.*, 1985; Bodmer *et al.*, 1987; Simpson *et al.*, 1987). In addition, several single gene disorders with unknown biochemical defects (e.g. Duchenne muscular dystrophy) have been mapped and candidate genes identified (Monaco *et al.*, 1986).

Efforts have begun to focus on disorders which 1) are common, 2) may show only partial family clustering, 3) are of unknown biochemical origins, 4) have variable, generally late age of onset, 5) are related to environmental or lifestyle conditions, 6) are of heterogeneous nature, and

7) may be polygenic. These include diabetes, epilepsy, hypertension, hyperlipoproteinemia, cardiovascular disease, and a range of psychiatric disorders (Schull and Weiss, 1980).

In these complex and multifactorial disorders, where disease occurrence depends upon the presence and the biologic interactions among genetic and environmental factors, susceptibility genes may be neither necessary, nor sufficient, for the expression of the disorder. A range of genetic factors may interact with numerous other contributing factors, leading to the differential expression of symptoms. It is often difficult in these cases to classify family members as affected, unaffected, exposed, or unexposed. Spurious associations can occur when families are drawn from subpopulations that are genetically heterogeneous at the marker locus, or where differences in the disease prevalence occur among population subgroups (Khoury *et al.*, 1990).

In an attempt to minimize these confounding effects in analysis, certain criteria must be placed on the selection of the study population. The populations should be more or less totally ascertainable and have minimal ethnic and lifestyle heterogeneity (Schull and Weiss, 1980). The Micronesian population of Nauru meets these criteria, although even in this relatively homogeneous population, there is evidence to suggest genetic diversity in diabetes (Serjeantson *et al.*, 1989).

The detection of associations between candidate genes and disease states at a population level, relies upon the presence of linkage disequilibrium between the polymorphic restriction site and the disease promoting genetic lesion. For large genes in regions with a high rate of recombination, this assumption may not hold. For instance, some insulin-receptor (INSR) RFLPs are not in linkage disequilibrium with each other (Elbein *et al.*, 1986; Serjeantson *et al.*, 1987), so that a single RFLP at INSR will not necessarily be associated with an INSR gene defect in population studies. Population studies can detect linkage up to about 100kb (≃0.1 cM) whereas family studies are capable of detecting linkage of up to 15 000kb (Lander and Botstein, 1986a). Moreover genetic heterogeneity can only be detected using linkage analysis.

5.2.2 Lipoproteinemias - family studies

It has been recognized that the hyperlipoproteinemias comprise a heterogeneous group of disorders, and these disorders have subsequently been subdivided into types I through V (Table 1.4). Genetic defects in the lipoprotein lipase and apolipoprotein C2 genes are known to lead to type I hyperlipoproteinemia (Humphries *et al.*, 1984; Langlois *et al.*, 1989; Monsalve *et al.*, 1990), whilst apolipoprotein E and low density lipoprotein receptor gene defects are causal factors in types III and IIa hyperlipoproteinemia respectively (Utermann *et al.*, 1979; Breslow *et al.*, 1982a; Leppert *et al.*, 1986; Brink *et al.*, 1990). Familial combined hyperlipoproteinemia (types IIb

and IV hyperlipoproteinemia) has recently been found to be linked to the APOA1-APOC3-APOA4 gene cluster (Wojciechowski *et al.*, 1991).

However, these hyperlipoproteinemia subtypes are neither genetically nor clinically homogeneous, and a great deal remains to be understood concerning possible genetic susceptibility to the hyperlipoproteinemias.

Whilst it may be clinically appropriate to attempt to divide the hyperlipoproteinemias into meaningful subtypes, and given that the recognised subtypes are genetically heterogeneous, it may be worthwhile to ignore *a priori* classifications and "classical diagnoses" and examine hypercholesterolemia and hypertriglyceridemia in the population in general rather than in predefined disease groups. This approach will be taken in the present chapter in a linkage analysis of the lipoprotein genes and hypercholesterolemia or hypertriglyceridemia, where these "affected states" are defined as the possession of plasma cholesterol or triglycerides levels above the 90th percentile.

5.2.3 Non-insulin dependent diabetes mellitus - family studies

Formal pedigree analysis of NIDDM in Caucasoid populations has been difficult due to the heterogeneity of environmental factors to which individuals in these populations are exposed (Zimmet, 1982).

The populations of the Seminole Amerindians, Pima Amerindians and Nauruan Micronesians have a high prevalence of NIDDM and a relatively homogeneous environment. In these populations pedigree analysis has provided evidence of familial clustering in NIDDM, and suggests the presence of a major gene in the development of the disease (Elston *et al.*, 1974; Serjeantson and Zimmet, 1984; Knowler *et al.*, 1990). Formal segregation analysis in Micronesians favours a dominant, but is compatible with a codominant, mode of inheritance for hyperglycaemia (Serjeantson and Zimmet, 1989).

There are as yet no known RFLP markers for NIDDM. Many genes have been examined, including the insulin gene, the insulin receptor gene, and the glucose transporter gene family, with claims and counter-claims of associations being reported in the literature (Serjeantson and Zimmet, 1991). Most recently Bell *et al.* (1991) have reported a maximum lod score for linkage between maturity-onset diabetes of the young (MODY) and the adenosine deaminase (ADA) gene of chromosome 20q. There is a rather early age of onset in NIDDM among Nauruans, which suggests that some MODY families may be part of the NIDDM spectrum seen in this population (Serjeantson and Zimmet, 1991). The ADA-related polymorphism was examined in the Nauruan population. The same two alleles, out of a possible five, were segregating in most of the families, and the lod scores were consequently uninformative. However, there was no distortion in allele frequencies in NIDDM patients of

early onset when compared with older, unaffected controls (pers. comm. Prof. S.W. Serjeantson).

The confusion in the literature may have arisen as a result of a dependence on population studies. If NIDDM is genetically heterogeneous, or has a polygenic aetiology then pedigree analyses of candidate genes, such as the lipoprotein genes are essential (Serjeantson and Zimmet, 1991).

5.3 MATERIALS AND METHODS

5.3.1 Pedigree selection and structure

Families were selected for linkage analysis of the lipoprotein genes and hyperlipoproteinemia on the basis of having at least two first degree relatives with plasma cholesterol or plasma triglycerides levels over the 90th percentile for the population, at the time of either the 1982 or 1987 survey. Eleven families were found to be segregating for high cholesterol levels, and nine for high triglyceride levels. Blood samples were available for 51 adults in the high cholesterol families and 44 adults in the high triglyceride families. Of the individuals in the pedigrees, 26 and 20 respectively, had cholesterol or triglyceride levels in the top decile.

Families which were segregating for NIDDM were selected from the 1982 Nauru Survey. Twelve families were selected and the diabetic status and age of the adult individuals examined. As the onset of NIDDM is age dependent, an age adjusted liability score based on measures of fasting and

2hr plasma glucose levels was calculated for each individual (Serjeantson and Zimmet, 1984). This liability score can be viewed as an estimate of the likelihood of progression to diabetes in younger individuals and in effect dichotomizes the population so that in every age group there is approximately the same proportion of individuals who are "liable to diabetes". A liability score of intermediate value indicates that the likelihood of progression to diabetes can not be determined.

Of the 60 individuals in the selected families 14 were diabetic. After applying the liability score all 14 were classified as liable to diabetes and were therefore classified as affected. Forty six individuals in the families were not diabetic; of these, two went on to develop diabetes by 1987 and were consequently reclassified as affected; ten had liability scores of intermediate value and were classified as being of unknown diabetic status; twenty eight had low liability scores and remained classified as unaffected. Six of the forty six individuals had high liability scores, which would indicate a tendency towards the later development of diabetes. A conservative decision was made to classify five of these six individuals as unknown, despite their high liability. The remaining individual was found to have impaired glucose tolerance in both 1982 and 1987, and was classified as affected.

The reassignment of "non-diabetic" to "unknown" on the basis of information gained by use of the liability scores resulted in the loss of two families from the sample. These

families were no longer informative for NIDDM after reassignment. The remaining ten families comprised 53 adults for whom there were blood samples. Of the family members, 17 were classified as affected, 10 as unknown, and 26 as unaffected.

The samples from both the hyperlipoproteinemia and the NIDDM pedigrees were typed for lipoprotein fragment length polymorphisms (Sections 5.3.2 and 5.3.3). Difficulties were experienced in obtaining results using the pLDLR-2HHI probe, and successful typing of the LDLR *Pvu*II site was limited to half of the sample.

5.3.2 APOA1, APOA2, APOC1, APOC2, APOD, LDLR and LPL RFLP typing

Genomic DNA was prepared from the samples, digested with *Dra*I, *Msp*I, *Pvu*II and *Taq*I and transferred to nylon membranes as described in Chapter 2. The cut DNA was hybridized to plasmids containing the cDNA probes discussed in Section 2.2.4 (refer also Table 3.7).

5.3.3 APOE AFLP typing

A 234bp fragment of APOE exon 4 was amplified from each sample and digested with *Cfo*I according to the methods described in Chapter 2.

5.3.4 Linkage analysis

Linkage analyses were performed using the LIPED computer programme (Ott, 1974), as modified by Dr A. Spence, and lod scores calculated (where lod score = $\log_{10}$(probability of

observed offspring at given rate of recombination / probability of observed offspring in absence of linkage)). All analyses were originally undertaken over a range of $\theta_f = \theta_m$ from 0.00 to 0.40 (where θ_f = female recombination rate, and θ_m = male recombination rate). As recombination rates are known to differ between the sexes (White and Lalouel, 1987), lod scores of particular interest were recalculated over ranges of $\theta_f \neq \theta_m$ from 0.05 to 0.50.

The criteria for the interpretation of lod scores set by Morton (1955), and based on a type I error of $\alpha = 0.001$, and a type II error of $\beta = 0.01$, were applied. With a lod score of 3 or greater, it is concluded that the frequency of recombination is significantly less than 0.50. If the lod score is -2 or less, it is concluded that the frequency of recombination is significantly greater than the value of θ for which the lods were calculated. At lod scores between 3 and -2 judgement of the likelihood of linkage is not possible as the pedigrees do not contain enough information.

These criteria are based upon the theorem that the expected value of a probability ratio is one under the null hypothesis, regardless of the alternative hypothesis (Haldane and Smith, 1947; Wald, 1947). Some workers have felt that this theorem is true only if the theorem is specified *a priori*, which is not the case, as in practice it is based upon the sample (Chotai, 1984). However, recently Collins and Morton (1991) have used exact

calculation to shown that posterior specification of the alternative hypothesis does not bias the test.

Lod scores were calculated under models of both dominant and recessive inheritance. Under a model of reduced penetrance it was necessary to assume reduced penetrance, as one of the hypercholesterolemic and hypertriglyceridemic pedigrees included affected offspring from unaffected parents. Although the same situation did not occur in the pedigrees used in the diabetes linkage analysis, the same assumption of reduced penetrance was employed, as there are cases in the population where non-diabetic parents have diabetic offspring.

Allele frequencies for the lipoprotein alleles were taken as those established in Chapter 3 (Table 3.9). Allele frequencies for hypercholesterolemia and hypertriglyceridemia were calculated on the basis of 10% of the population being affected and were as follows:- under a dominant model, $p = 0.051$, $q = 0.949$; under a recessive model, $p = 0.316$, $q = 0.684$; where p is the frequency of the "affected" allele and q is the frequency of the "normal" allele. The allele frequencies for NIDDM in the population were estimated based on the proportion of over-forty-year-olds who had diabetes (46.9%), as no individuals in this age group had a liability score which classified them as being of unknown diabetic status. The estimated allele frequencies were: under a dominant model, $p = 0.271$, $q = 0.729$; under a recessive model, $p = 0.685$, $q = 0.315$;

where p is the frequency of the "affected" allele and q is the frequency of the "normal" allele.

5.4 RESULTS

5.4.1 Micronesian lipoprotein genes in hypercholesterolemia - linkage analysis

The eleven pedigrees used in the analysis of linkage between the eight lipoprotein fragment length polymorphisms and hypercholesterolemia are presented in Table 5.1. Lod scores were calculated under models of recessive inheritance with full penetrance, and dominant inheritance with 99%, 90% and 80% penetrance, at values of $\theta_f = \theta_m$ from 0.0 to 0.5 (Table 5.2). Under a dominant model of inheritance close linkage of hypercholesterolemia was excluded for APOA1, APOC1, APOC2, APOD, APOE and LPL at 99% penetrance. Under a recessive model linkage of APOD was excluded, while lod scores were uninformative for the remaining loci. Lod scores were uninformative for LDLR, under both models of inheritance, and at all levels of penetrance tested. A small, but positive lod score of 0.97 was obtained at $\theta_f = \theta_m = 0.1$ for APOA2.

Under a dominant model with 99% penetrance the negative lod scores for APOC1, APOC2, APOD, APOE, LDLR and LPL were due largely to the influences of the minority of families, with between 60% and 96% of the magnitude of the lod scores being attributable to one or two families.

Further lod scores for APOA2 and hypercholesterolemia were calculated at values where $\theta_f \neq \theta_m$ to establish the effect

of different recombination rates for females and males (Table 5.3). Maximum, but uninformative lod scores for APOA2 of 1.11 were obtained at θ_f = 0.35 to 0.50 and θ_m = 0.05.

When the maximum APOA2, hypercholesterolemia lod scores were examined for the contribution of individual families (Table 5.4), three families were found to be largely responsible for the positive lod scores, with a combined lod score across these families of 1.61 and 1.55. Five families had lod scores extremely close to zero (-0.05 to 0.05), whilst the remaining three had lod scores ranging from -0.07 to -0.25.

Four of the unaffected individuals in the high cholesterol pedigrees were found to have plasma cholesterol levels above the 75th percentile of 5.86 mM, which had been used to define high and low groups in the population in Section 4 (Refer Table 5.1). When the status of these individuals was reassigned to unknown, and the linkage analysis repeated, the resultant lod scores precipitated an altered conclusion from the original analyses, under dominant inheritance with reduced penetrance, in only one situation. The lod score for linkage of the LDLR *Pvu*II RFLP with hypercholesterolemia, following reassignment of the four individuals was -2.81 at $\theta_f = \theta_m = 0.0$, compared with an original lod score of -1.59 (Refer Table 5.2). The lod scores for the remaining seven loci differed from the original lod scores by between 0.01 and 1.69, with the APOA2 analysis remaining uninformative, and results from

the other six loci still providing evidence of exclusion to linkage with hypercholesterolemia.

Table 5.1 Extreme plasma cholesterol pedigrees.

Pedigree	Ind.†	Status§	Locus							
			A1	A2	C1	C2	D	E	LDLR	LPL
A										
	Mo	2	AB	AA	AB	AB	BB	AB	BB	AA
	C1	1	AB	AA	AB	AA	AB	BB		AB
	C2	2	AA	AA	AA	BB	AB	AB	BB	AA
	C3	2	AB	AA	AA	AB	AB	AB	AB	AB
	C4	1	AB	AA	AB	BB	AB	BB		AA
B										
	C1	2	AB	AB	AA	AA	AB	BB	BB	AB
	C2	1	AB	AA	AA	AB	AB	BB		AA
	C3	1	AB	AA	AA	AB	AB	BB		AA
C										
	C1	2	AB	AA		AB		BB		
	C2	2	AA	AA	AB	AB	AB	BB	BB	AA
	C3	1	AA	BB	AA	BB	AA	BC		AA
	C4	2	AA	AA	AB	BB	AA	BC	BB	AA
	C5	1	AA	AB	AB	AA	AB	BB		AA
	C6	1	AB	AB	AA	BB	AB	BC		AA
D										
	Mo	2	AB	AA	AA	AB	BB	BB	BB	AB
	C1	1	AA	AB	AA	AB	BB	BB		BB
	C2	2	AA	AA	AB	AB	BB	BC	AB	AB
	C3	1			AB	AB	BB	BC		
	C4	2#	AA	AA	AB	AB	BB	BC	AB	AA
	C5	2	AB	AA	AA	AB	AB	BB		AA
E										
	Mo	2#	AA	AB	AB	AB	BB	BB	AB	AB
	Fa	2	AB	AB	AA	BB	BB	BB	BB	AA
	C1	2	AB	AB	AA	BB	BB	BB	AB	AA
	C2	1	AA	AA	AA	BB	BB	BB		AA
	C3	2	AB	AA	AB	AB	BB	BB	BB	AA
	C4	1	AB	AB	AB	AB	BB	BB		AB
	C5	1	AA	AB	AA	BB	BB	BB		AB
	C6	2	AA	AB	AA	BB	BB	BB		AA

† Ind. Individuals:- Mo = mother; Fa = father; Cn = numbered children.

§ Status 1 = affected; status 2 = unaffected.

Continued overleaf.

Plasma cholesterol level above the 75th percentile.

Table 5.1 Cont'd. Extreme plasma cholesterol pedigrees.

Pedigree	Ind.†	Status§	Locus							
			A1	A2	C1	C2	D	E	LDLR	LPL
F										
	C1	2	AB	AB	AA	AB	BB	BB		AA
	C2	1	AB	AB	AA	AA	AB	BB		AB
	C3	1	BB	AA	AA	AA	AB	BB	BB	AA
G										
	C1	1	AB	BB	AA	BB	AB	BC		AB
	C2	2	AB	BB	AA	AB	BB	BC	AB	AA
	C3	1	AB	AB	AA	AB	BB	BC		AA
	C4	1	AA	BB	AA	AB	BB	BC		AA
	C5	2	AA	AB	AA	BB	BB	BC		AB
H										
	Fa	2	AA	AA	AA	AB		BB		
	C1	1	AB	AB	AA	AB	BB	BC		AB
	C2	1	AB	AB	AA	AB	AB	BC		AB
I										
	C1	2#	AA	AA		BB	AB	BB	AB	AA
	C2	1			AB	AB	AB	BC		AA
	C3	2#						BC		
	C4	1			AB	BB	BB	BB		AA
	C5	2	AA	AA	BB	AA	BB	BC		AA
J										
	C1	1	AA	AB	AA	BB	AB	BB		AB
	C2	2	AB	AB	AA	AB	AB	BB	AB	AA
	C3	2	AA	AB		BB	AB	BB	BB	AB
	C4	1	AA	AB	AA	BB	BB	BB		AB
K										
	Mo	2	AB	AB	AA	AB	BB	BC		AB
	C1	1	AA		AA	AB	BB	BC		AA
	C2	1	BB	AB	AB	BB	BB	CC		AA

† Ind. Individuals:- Mo = mother; Fa = father; Cn = numbered children.
§ Status 1 = affected; status 2 = unaffected.
# Plasma cholesterol level above the 75th percentile.

Table 5.2 Lod scores for linkage of hypercholesterolemia with lipoprotein gene fragment length polymorphisms.

Locus	$\theta_f = \theta_m$ †	Dominant Penetrance			Recessive Penetrance
		0.99	0.90	0.80	1.00
APOA1	0.4	-0.07	-0.06	-0.06	-0.09
(*Msp*I)	0.3	-0.31	-0.27	-0.23	-0.39
	0.2	-0.73	-0.64	-0.55	-0.97
	0.1	-1.40	-1.23	-1.07	-2.04
	0.0	-2.61	-2.29	-2.06	-1.89
APOA2	0.4	0.17	0.13	0.09	0.08
(*Msp*I)	0.3	0.52	0.41	0.32	0.23
	0.2	0.84	0.70	0.56	0.32
	0.1	0.97	0.88	0.75	0.15
	0.0	0.02	0.76	0.79	0.37
APOC1	0.4	-0.05	-0.04	-0.03	0.02
(*Dra*I)	0.3	-0.23	-0.17	-0.12	0.03
	0.2	-0.60	-0.43	-0.33	-0.06
	0.1	-1.33	-0.92	-0.70	-0.44
	0.0	-3.31	-1.95	-1.46	-0.49
APOC2	0.4	-0.07	-0.06	-0.05	-0.08
(*Taq*I)	0.3	-0.33	-0.27	-0.22	-0.37
	0.2	-0.85	-0.69	-0.57	-1.04
	0.1	-1.88	-1.48	-1.21	-2.59
	0.0	-8.23	-4.22	-3.08	-0.38
APOD	0.4	-0.08	-0.06	-0.05	-0.05
(*Taq*I)	0.3	-0.33	-0.25	-0.19	-0.22
	0.2	-0.80	-0.60	-0.46	-0.56
	0.1	-1.61	-1.13	-0.85	-1.23
	0.0	-3.19	-1.85	-1.37	-2.70
APOE	0.4	-0.05	-0.04	-0.02	0.01
(*Cfo*I)	0.3	-0.23	-0.15	-0.09	0.01
	0.2	-0.58	-0.38	-0.24	-0.12
	0.1	-1.22	-0.76	-0.51	-0.53
	0.0	-2.50	-1.35	-0.97	-1.60
LDLR	0.4	-0.02	-0.01	-0.01	-0.01
(*Pvu*II)	0.3	-0.09	-0.06	-0.04	-0.02
	0.2	-0.23	-0.14	-0.08	-0.05
	0.1	-0.53	-0.29	-0.16	-0.09
	0.0	-1.59	-0.56	-0.28	-0.16
LPL	0.4	0.05	0.05	0.04	0.05
(*Pvu*II)	0.3	0.14	0.13	0.11	0.16
	0.2	0.08	0.13	0.14	0.21
	0.1	-0.37	-0.10	-0.00	0.01
	0.0	-2.07	-0.95	-0.62	-0.19

† θ_f = female recombination fraction; θ_m = male recombination fraction.

Table 5.3 Lod scores for linkage of hypercholesterolemia with APOA2 *Msp*I restriction fragment length polymorphism under a model of dominant inheritance and 99% penetrance.

Locus		θ_f†			
	θ_m†	0.05	0.20	0.35	0.50
APOA2 (*Msp*I)	0.05	0.88	1.07	1.11	1.11
	0.20	0.86	0.84	0.76	0.72
	0.35	0.63	0.50	0.34	0.26
	0.50	0.47	0.30	0.10	0.00

† θ_f = female recombination fraction; θ_m = male recombination fraction.

Table 5.4 Family lod scores for linkage of hypercholesterolemia with APOA2 *Msp*I restriction fragment length polymorphism under a model of dominant inheritance with 99% penetrance.

Pedigree	θ_f†	
	0.35	0.50
D	0.80	0.80
C	0.60	0.56
B	0.21	0.19
D, C, and B	1.61	1.55
I	0.03	0.03
H	0.02	0.00
K	0.00	0.00
A	-0.02	-0.02
J	-0.03	-0.02
I, H, K, A, and J	0.00	0.01
F	-0.08	-0.07
G	-0.17	-0.14
E	-0.25	-0.22
F, G, and E	-0.50	-0.43
Total	1.11	1.11

† θ_f = female recombination fraction; θ_m = male recombination fraction = 0.05.

5.4.2 Micronesian lipoprotein genes in hypertriglyceridemia - linkage analysis

Pedigrees used in the linkage analysis of the eight lipoprotein fragment length polymorphisms and hypertriglyceridemia are presented in Table 5.5. Lod scores were calculated under models of recessive inheritance with full penetrance, and dominant inheritance with 99%, 90% and 80% penetrance, at values of $\theta_f = \theta_m$ from 0.0 to 0.5 (Table 5.6). Under a recessive model linkage of APOC2 was excluded, while lod scores were uninformative for the remaining loci. Under a dominant model of inheritance close linkage of hypertriglyceridemia was excluded for APOA1, APOA2, APOC2, and APOD at 99% penetrance. Under the same model, lod scores were uninformative for APOC1, APOE, LDLR and LPL.

In the case of APOC1 under dominant inheritance, an examination of lod scores for individual families revealed that one family had a positive lod score of 1.19 at $\theta_f = \theta_m = 0.0$, and the remaining families had lod scores which were negative or close to zero. For APOA2, APOC2, and APOD, under dominant penetrance of 99%, one or two families were responsible for over 75% of the magnitude of the negative lod scores.

Five of the unaffected individuals in the high triglyceride pedigrees were found to have plasma triglyceride levels above the 75th percentile of 1.59 mM, which had been used to define high and low groups in the population in Section 4 (Refer Table 5.5). When the status of these individuals was reassigned to unknown, and the linkage analysis

repeated, the resultant lod scores, under a model of dominant inheritance with 99% penetrance, and at $H_f = \theta_m = 0.0$, differed by 0.01 to 1.19 from those of the original analysis. However, the change in lod scores, with the reclassification of the five individuals, did not alter the nature of the conclusions drawn: the APOC1, APOE, LDLR and LPL fragment length polymorphisms remained uninformative, while the lod scores for the APOA1, APOA2, APOC2 and APOD RFLPs remained less than -2, and so provided evidence for the exclusion of these loci from linkage with hypertriglyceridemia.

Table 5.5 Extreme plasma triglyceride pedigrees.

Pedigree	Ind.†	Status§	Locus							
			A1	A2	C1	C2	D	E	LDLR	LPL
L										
	C1	1	AA	AA	AA	BB	AB	BC		AB
	C2	2	AB	AB	AA	BB	AB	BC	BB	AA
	C3	1	AA	AB	AA	AB	AB	BB		AA
	C4	1	AB	AB	AA	AB	AB	BB		AA
M										
	C1	1	AB	AB	AA	AB	BB	BB		AA
	C2	1	BB	AA	AA	BB	BB	BB		AA
	C3	2	BB	AB	AA	BB	BB	BB	BB	AB
E										
	Mo	2	AA	AB	AB	AB	BB	BB	AB	AB
	Fa	2	AB	AB	AA	BB	BB	BB	BB	AA
	C1	2	AB	AB	AA	BB	BB	BB	AB	AA
	C2	1	AA	AA	AA	BB	BB	BB		AA
	C3	1	AB	AA	AB	AB	BB	BB	BB	AA
	C4	2	AB	AB	AB	AB	BB	BB		AB
	C5	2#	AA	AB	AA	BB	BB	BB		AB
	C6	2	AA	AB	AA	BB	BB	BB		AA
N										
	C1	1			AA			AB		AA
	C2	1	BB	AB	AA	AB	AB	AC		AA
	C3	1	AB	AB	AA	AB	AB	AC	AB	AA
	C4	2	AB	AB	AA	AA	AB	BB	AB	AA
O										
	C1	1	AA	BB	AA	AB	BB	BB		AB
	C2	2	AA	BB	AA	AB	BB	BB	BB	AB
	C3	1	AA	AB	AA	AB	AB	BB		AB
	C4	2#	AB	AA	AA	BB	BB	BB	BB	BB
	C5	2	AB	BB	AA	BB	BB	BB	BB	AA

† Ind. Individuals:- Mo = mother; Fa = father;
Cn = numbered children.
§ Status 1 = affected; status 2 = unaffected.
Continued overleaf
Plasma triglyceride level above the 75th percentile.

Table 5.5 Cont'd. Extreme plasma triglyceride pedigrees.

Pedigree	Ind.†	Status§	Locus							
			A1	A2	C1	C2	D	E	LDLR	LPL
P										
	C1	2#	AB	AA	AB	AB	BB	BB	BB	AB
	C2	1	BB	AA	AA	AB	BB	BB		AA
	C3	1	BB	AA	AA	AA	BB	BB		
	C4	2	BB	AA	AA	AB	BB	BB	BB	AB
Q										
	Mo	2	AB	AB	AA	AA	BB	AB	AB	AA
	C1	2		AA		AB		BB		
	C2	1	AA	AB	AA	AB	AB	BB	AB	AA
	C3	1	AB	AB	AA	AA	BB	AB		AA
R										
	Fa	2	AB	AA	AA	AA	BB	BB	AB	AB
	C1	2	AA	AA		AB	BB	BC	BB	BB
	C2	2	AB	AA	AB	AA	BB	BB	AB	BB
	C3	1	AB	AA	AA	AA	BB	BB		BB
	C4	2	AA	AA	AB	AA	BB	BB	AB	AB
	C5	2#	AB	AA	AB	AA	BB	BB	AB	AB
	C6	2#	AB	AA		AA	BB	BB	BB	BB
	C7	2	AB	AA		AA	BB	BB		AB
	C8	1	AB	AA	AA	AA	BB	BB		BB
K										
	Mo	2	AB	AB	AA	AB	BB	BC		AB
	C1	1	AA		AA	AB	BB	BC		AA
	C2	1	BB	AB	AB	BB	BB	CC		AA

† Ind. Individuals:- Mo = mother; Fa = father; Cn = numbered children.
§ Status 1 = affected; status 2 = unaffected.
# Plasma triglyceride level above the 75th percentile.

Table 5.6 Lod scores for linkage of hypertriglyceridemia with lipoprotein gene fragment length polymorphisms.

Locus	$\theta_f = \theta_m$ †	Dominant Penetrance			Recessive Penetrance
		0.99	0.90	0.80	1.00
APOA1	0.4	-0.04	-0.04	-0.03	-0.06
(*Msp*I)	0.3	-0.17	-0.15	-0.13	-0.25
	0.2	-0.45	-0.39	-0.34	-0.70
	0.1	-0.98	-0.82	-0.71	-1.72
	0.0	-2.34	-1.80	-1.54	-1.75
APOA2	0.4	-0.04	-0.02	-0.01	0.01
(*Msp*I)	0.3	-0.18	-0.12	-0.07	0.07
	0.2	-0.46	-0.32	-0.19	0.16
	0.1	-1.01	-0.64	-0.40	0.16
	0.0	-3.33	-1.32	-0.75	0.65
APOC1	0.4	0.12	0.10	0.07	0.04
(*Dra*I)	0.3	0.31	0.26	0.20	0.09
	0.2	0.39	0.35	0.30	0.12
	0.1	0.27	0.30	0.39	-0.04
	0.0	-0.38	-0.05	0.07	0.42
APOC2	0.4	-0.05	-0.07	-0.05	-0.05
(*Taq*I)	0.3	-0.23	-0.28	-0.20	-0.20
	0.2	-0.69	-0.74	-0.52	-0.58
	0.1	-1.82	-1.62	-1.12	-1.42
	0.0	-4.85	-3.08	-2.08	-2.56
APOD	0.4	-0.01	0.00	-0.01	-0.02
(*Taq*I)	0.3	-0.05	-0.05	-0.06	-0.06
	0.2	-0.16	-0.16	-0.15	-0.16
	0.1	-0.50	-0.45	-0.37	-0.40
	0.0	-2.29	-1.26	-0.94	-1.18
APOE	0.4	0.00	0.00	0.00	0.00
(*Cfo*I)	0.3	-0.02	-0.04	-0.04	-0.02
	0.2	-0.19	-0.17	-0.14	-0.17
	0.1	-0.72	-0.47	-0.34	-0.62
	0.0	-1.69	-0.79	-1.56	-0.75
LDLR	0.4	0.01	0.00	0.00	0.00
(*Pvu*II)	0.3	0.04	0.03	0.02	0.00
	0.2	0.09	0.06	0.03	0.00
	0.1	0.16	0.10	0.06	-0.02
	0.0	0.22	0.15	0.09	-0.02
LPL	0.4	0.01	0.02	0.01	0.05
(*Pvu*II)	0.3	-0.03	0.00	0.03	0.18
	0.2	-0.12	-0.03	0.04	0.39
	0.1	-0.37	-0.15	0.03	0.66
	0.0	-1.30	-0.35	0.05	1.01

† θ_f = female recombination fraction; θ_m = male recombination fraction.

5.4.3 Micronesian lipoprotein genes in non-insulin dependent diabetes mellitus - linkage analysis

The pedigrees used in the analysis of linkage of the eight lipoprotein genes with diabetic status are presented in Table 5.7. Lod scores were calculated under models of recessive inheritance with full penetrance, and dominant inheritance with 99%, 90% and 80% penetrance, at values of $\theta_f = \theta_m$ from 0.0 to 0.5 (Table 5.8). Lod scores were uninformative under a model of recessive inheritance. When dominance penetrance of 99% was modelled, close linkage could be excluded for APOA1, APOC2, APOD and APOE. The lod scores for APOA2, APOC1, LDLR and LPL were uninformative under models of dominant inheritance.

The exclusion of linkage under a dominant model with 99% penetrance for APOA1, APOC2, APOD and APOE was mainly due to the contribution of lod scores from one or two families, with these lod scores accounting for from 65% to 82% of the magnitude of the combined lod scores.

The lod scores for APOD were also calculated over a range of $\theta_f \neq \theta_m$ from 0.05 to 0.50, assuming dominant inheritance with 99% penetrance. However, the magnitudes of the lod scores from this analysis were no greater than those calculated when $\theta_f = \theta_m$.

The positive, but uninformative lod score of 1.50 at $\theta_f = \theta_m = 0.0$, for APOA2, calculated under assumptions of dominant and recessive inheritance, with 99% and full penetrance respectively, prompted further investigation. The lod scores for APOA2 were recalculated over a range of

$\theta_f \neq \theta_m$ from 0.05 to 0.50. None of the resultant lod scores were larger than those calculated under assumptions of equal recombination rates for females and males. The lod scores of individual families were examined (Table 5.9), and under a recessive model four families were found to be the major contributors to a positive lod score of 1.87. Under a dominant model three families accounted for the majority of a positive lod score of 1.70.

Linkage analysis was also performed modelling for a gene dosage effect. In this situation the likelihood of diabetes, given two "affected" alleles was 0.99, and given one "affected" allele was 0.50. The corresponding allele frequencies were $p = 0.469$, $q = 0.531$; where p is the frequency of the "affected" allele and q is the frequency of the "normal allele". Under the assumption of a gene dosage effect, lod scores at $\theta_f = \theta_m = 0.0$, ranged from -0.50 to 0.49, and consequently provided insufficient evidence to accept or exclude linkage of these markers with diabetic liability.

Table 5.7 Diabetic liability pedigrees.

Pedigree	Ind.†	Status§	Locus							
			A1	A2	C1	C2	D	E	LDLR	LPL
A										
	Mo	1	AB	AA	AB	AB	BB	AB	BB	AA
	C1	1	AB	AA	AB	AA	AB	BB		AB
	C2		AA	AA	AA	BB	AB	AB	BB	AA
	C3		AB	AA	AA	AB	AB	AB	AB	AB
	C4	2	AB	AA	AB	BB	AB	BB		AA
B										
	C1		AB	AB	AA	AA	AB	BB	BB	AB
	C2	2	AB	AA	AA	AB	AB	BB		AA
	C3	1	AB	AA	AA	AB	AB	BB		AA
C										
	C1		AB	AA		AB		BB		
	C2	2	AA	AA	AB	AB	AB	BB	BB	AA
	C3	1	AA	BB	AA	BB	AA	BC		AA
	C4		AA	AA	AB	BB	AA	BC	BB	AA
	C5	1	AA	AB	AB	AA	AB	BB		AA
	C6	2	AB	AB	AA	BB	AB	BC		AA
D										
	Mo	1	AB	AA	AA	AB	BB	BB	BB	AB
	C1	1	AA	AB	AA	AB	BB	BB		BB
	C2	2	AA	AA	AB	AB	BB	BC	AB	AB
	C3				AB	AB	BB	BC		
	C4	2	AA	AA	AB	AB	BB	BC	AB	AA
	C5	2	AB	AA	AA	AB	AB	BB		AA
E										
	Mo	2	AA	AB	AB	AB	BB	BB	AB	AB
	Fa	1	AB	AB	AA	BB	BB	BB	BB	AA
	C1	1	AB	AB	AA	BB	BB	BB	AB	AA
	C2	2	AA	AA	AA	BB	BB	BB		AA
	C3	2	AB	AA	AB	AB	BB	BB	BB	AA
	C4	2	AB	AB	AB	AB	BB	BB		AB
	C5	2	AA	AB	AA	BB	BB	BB		AB
	C6	2	AA	AB	AA	BB	BB	BB		AA

† Ind. Individuals:- Mo = mother; Fa = father; Cn = numbered children.

§ Status 1 = affected; status 2 = unaffected (not liable); blank = unknown.

Continued overleaf.

Table 5.7 Cont'd. Diabetic liability pedigrees.

Pedigree	Ind.†	Status§	Locus							
			A1	A2	C1	C2	D	E	LDLR	LPL
I										
	C1	2	AA	AA		BB	AB	BB	AB	AA
	C2	2			AB	AB	AB	BC		AA
	C3							BC		
	C4	1			AB	BB	BB	BB		AA
	C5	2	AA	AA	BB	AA	BB	BC		AA
K										
	Mo	2	AB	AB	AA	AB	BB	BC		AB
	C1	1	AA		AA	AB	BB	BC		AA
	C2	2	BB	AB	AB	BB	BB	CC		AA
N										
	C1	1			AA			AB		AA
	C2	2	BB	AB	AA	AB	AB	AC		AA
	C3	1	AB	AB	AA	AB	AB	AC	AB	AA
	C4	2	AB	AB	AA	AA	AB	BB	AB	AA
Q										
	Mo	2	AB	AB	AA	AA	BB	AB	AB	AA
	C1	2		AA		AB		BB		
	C2	1	AA	AB	AA	AB	AB	BB	AB	AA
	C3	1	AB	AB	AA	AA	BB	AB		AA
R										
	Fa		AB	AA	AA	AA	BB	BB	AB	AB
	C1	1	AA	AA		AB	BB	BC	BB	BB
	C2		AB	AA	AB	AA	BB	BB	AB	BB
	C3		AB	AA	AA	AA	BB	BB		BB
	C4	2	AA	AA	AB	AA	BB	BB	AB	AB
	C5	2	AB	AA	AB	AA	BB	BB	AB	AB
	C6	2	AB	AA		AA	BB	BB	BB	BB
	C7	1	AB	AA		AA	BB	BB		AB
	C8	2	AB	AA	AA	AA	BB	BB		BB

† Ind. Individuals:- Mo = mother; Fa = father; Cn = numbered children.

§ Status 1 = affected; status 2 = unaffected (not liable); blank = unknown.

Table 5.8 Lod scores for linkage of diabetic liability with lipoprotein gene fragment length polymorphisms.

Locus	$\theta_f = \theta_m$ †	Dominant Penetrance			Recessive Penetrance
		0.99	0.90	0.80	1.00
APOA1	0.4	-0.05	-0.03	-0.02	-0.01
(*Msp*I)	0.3	-0.24	-0.16	-0.10	-0.07
	0.2	-0.68	-0.43	-0.26	-0.21
	0.1	-1.68	-0.98	-0.56	-0.54
	0.0	-5.76	-2.19	-1.13	-1.27
APOA2	0.4	0.13	0.11	0.08	0.15
(*Msp*I)	0.3	0.45	0.36	0.29	0.49
	0.2	0.81	0.68	0.55	0.88
	0.1	1.16	0.99	0.82	1.24
	0.0	1.50	1.30	1.10	1.50
APOC1	0.4	-0.03	-0.02	-0.01	-0.08
(*Dra*I)	0.3	-0.11	-0.08	-0.06	-0.34
	0.2	-0.27	-0.19	-0.13	-0.87
	0.1	-0.56	-0.38	-0.25	-1.97
	0.0	-1.03	-0.65	-0.43	-0.63
APOC2	0.4	-0.01	-0.01	-0.02	-0.06
(*Taq*I)	0.3	-0.08	-0.08	-0.09	-0.27
	0.2	-0.36	-0.30	-0.26	-0.80
	0.1	-1.15	-0.82	-0.64	-2.11
	0.0	-5.22	-2.26	-1.49	-0.51
APOD	0.4	-0.03	-0.02	-0.02	-0.05
(*Taq*I)	0.3	-0.13	-0.10	-0.08	-0.22
	0.2	-0.34	-0.25	-0.19	-0.58
	0.1	-0.75	-0.52	-0.38	-1.29
	0.0	-2.21	-1.05	-0.69	-0.83
APOE	0.4	-0.01	-0.01	-0.01	-0.03
(*Cfo*I)	0.3	-0.10	-0.07	-0.06	-0.15
	0.2	-0.37	-0.24	-0.17	-0.48
	0.1	-0.99	-0.60	-0.41	-1.33
	0.0	-2.83	-1.36	-0.88	-0.98
LDLR	0.4	-0.01	0.00	0.00	0.02
(*Pvu*II)	0.3	-0.03	-0.02	-0.01	0.08
	0.2	-0.08	-0.05	-0.02	0.16
	0.1	-0.19	-0.11	-0.05	0.21
	0.0	-0.45	-0.22	-0.11	0.19
LPL	0.4	0.02	0.01	0.01	-0.03
(*Pvu*II)	0.3	0.04	0.02	0.01	-0.18
	0.2	0.02	0.01	-0.01	-0.57
	0.1	-0.04	-0.06	-0.06	-1.48
	0.0	-0.16	-0.18	-0.17	-0.62

† θ_f = female recombination fraction; θ_m = male recombination fraction.

Table 5.9 Family lod scores for linkage of diabetic liability with APOA2 *Msp*I restriction fragment length polymorphism.

Pedigree	Mode of inheritance	
	Dominant Penetrance 0.99	Recessive Penetrance 1.00
D	0.89	0.90
Q	0.44	0.47
E	0.29	0.30
C	0.08	0.20
D, Q, E, and C	1.70	1.87
I	0.03	0.05
K	0.00	0.00
R	0.00	0.00
I, K, and R	0.03	0.05
A	-0.02	-0.02
N	-0.07	-0.09
B	-0.14	-0.31
A, N, and B	-0.23	-0.42
Total	1.50	1.50

θ_f = female recombination fraction = 0.50; θ_m = male recombination fraction = 0.50.

5.5 DISCUSSION

5.5.1 Exclusion of linkage to hypercholesterolemia, hypertriglyceridemia, and diabetic liability

Linkage of diabetic liability, hypercholesterolemia and hypertriglyceridemia to a number of the marker loci tested has been excluded in Micronesians. The APOA1, APOC2, and APOD fragment length polymorphisms were excluded from linkage to all three disorders, when modelled under a dominant mode of inheritance. Under the same model APOE polymorphism was excluded from linkage with diabetic liability and hypercholesterolemia. The three diallelic markers, APOA1, APOC2 and APOD, had the highest polymorphic

information content (PIC) (Section 3.4.3, Table 3.9). Whilst the PIC for APOE was not as high, this marker was successfully typed for every individual, in every pedigree. This was not the case for the LDLR *Pvu*II RFLP, which was consequently uninformative in all but one of the analyses. Under a model of dominant inheritance the APOA2 locus has been excluded from linkage to hypertriglyceridemia, and the APOC1 and LPL loci excluded from linkage to hypercholesterolemia.

The two loci, APOC2 and APOD, which showed associations with plasma triglyceride levels and diabetes respectively (Section 4.4), where not implicated in these disorders in formal linkage analysis. Indeed APOC2 was excluded from linkage to hypertriglyceridemia under models of both dominant and recessive inheritance, and from linkage to diabetic liability, under dominant inheritance. The APOD locus was excluded from linkage to all three disorders, under dominant inheritance, though from hypercholesterolemia only under a recessive model of inheritance.

This highlights the problem of undertaking linkage analysis when the precise mode of inheritance is unknown. Segregation analysis for hypercholesterolemia and hypertriglyceridemia has not been undertaken in the Micronesian population. As a result it is not known whether the modes of inheritance for these disorders are indeed compatible with assumptions of either dominant inheritance,

with reduced penetrance, or recessive inheritance with full penetrance.

Segregation analysis of hypoglycaemia in the Micronesian population has been performed, and results favour a dominant mode of inheritance (Serjeantson and Zimmet, 1989). However, the results of the segregation analysis are also compatible with a codominant mode of inheritance. One possible expression of a codominant genotype in this situation could be a lowered chance of phenotypic expression of the disorder; that is a gene dosage effect, where the chance of diabetes given affected homozygosity at the disease locus, would be 0.99, and given heterozygosity, would be 0.50. Under such assumptions, the lod scores from the current analysis were uninformative. The possibility of linkage between the lipoprotein loci under investigation, and diabetic liability, inherited as a codominant disorder with an accompanying gene dosage effect has therefore not been excluded.

The APOC2 *Taq*I polymorphism was found to affect plasma triglyceride levels above and below the 75th percentile (Section 4.4.1.5) and when plasma triglyceride concentration was treated as a quantitative variable (Section 4.4.3.1). In the current analysis the reclassification of "unaffected" individuals to "unknown" on the basis of possession of plasma triglyceride levels over the 75th percentile did not alter the nature of the conclusions made on the basis of lod scores. Linkage of the

APOC2 RFLP with hypertriglyceridemia, was excluded under dominant inheritance with 99% penetrance.

The exclusion of linkage of the APOC2 and APOD RFLPs to hypertriglyceridemia and diabetic liability may indicate that the assumptions concerning mode of inheritance made in the linkage analyses were not valid.

5.5.2 Limitations of single marker linkage analysis

The use of single-markers may be inefficient and not a powerful enough technique to successfully establish linkage in complex traits (Lander and Botstein, 1986*b*). In particular, genetic heterogeneity can confound single-marker linkage analysis. Two sources of genetic heterogeneity may affect linkage analysis. Firstly, the term "genetic heterogeneity" will be used to refer to the situation where alterations in different genes may result in the same disorder; within one ethnic group, but most particularly between ethnic groups, the predisposing genetic defect or defects may be different. Secondly, the term "polygenic heterogeneity" will be used to refer to the situation where disorder may be polygenic, with a number of loci, or different combinations of subsets of that number, having an influence in setting the genetic predisposition to the disorder.

A locus of interest may be responsible for only a fraction of the occurrences of a heterogeneous trait. According to Lander and Botstein (1986*b*) the chance that the marker will fail to cosegregate with the disorder is the apparent

recombination fraction $\theta' = \theta\alpha + 1/2(1 - \alpha)$, where α is the fraction of occurrences of the disorder attributable to the locus of interest. If α is small, linkage will appear to be loose, even when θ is also small. Even if linkage is detected it is difficult to estimate the values of α and θ, as the effects of low α and high θ are the same.

When there is an *a priori* reason for expecting genetic heterogeneity, families may be subdivided on the basis of positive or negative lod scores. To establish whether the heterogeneity between the families is significant Morton (1956) applied the likelihood ratio test, in what is now referred to as the pre-divided sample test (Hodge *et al.*, 1983). However, this test relies on asymptotic theory for its interpretation, which is thought to be very imprecise for small families (White and Lalouel, 1987). The practice of subgrouping families according to lod score values, can greatly inflate the sum of the lods. It is not thought appropriate to emphasise the resultant higher positive lod scores in the current study, where the families are too small for application of the predivided sample test. In addition, there is no evidence for clinical heterogeneity in hypercholesterolemia, hypertriglyceridemia, or NIDDM in the Micronesian population.

Although only those individuals who claimed Nauruan ancestry were included in the samples, financial benefits can and have accrued from such claims, and it is not unreasonable to expect some respondents to have falsely stated their ethnicity. It has been noted that it is more

likely for Nauruan individuals of mixed ethnicity to intermarry, leading to a concentration of non-Micronesian genetic influences within particular pedigrees (pers. comm. Prof. S.W. Serjeantson).

The effect of foreign admixture on the occurrence of NIDDM in Nauru has been recognised (Serjeantson *et al.*, 1983). In over forty year olds, 83% of full blooded Nauruans, as determined by HLA typing, were diabetic, compared with only 17% of those with ancestral foreign admixture. In addition, the presence of HLA-Bw22(Bw56) was found to be associated with a small increased risk for early onset of diabetes. The increase of HLA-Bw56 frequency in diabetics was consistent with its increase in Micronesians in general. HLA-Bw56 is a marker for Nauruan ancestry not NIDDM (Serjeantson and Zimmet, 1989).

APOC2 and APOD, were implicated as genetic markers for hypertriglyceridemia and diabetes respectively (Section 4.4). The difference in APOC2 *Taq*I allele frequencies between Micronesians and Caucasoids, although significant was only 0.120 (Presence of site: Micronesians 0.530 ± 0.035; Pooled Caucasoids 0.41 ± 0.02). The APOD *Taq*I site, whilst being more frequent in Micronesians (0.363 ± 0.037), was present at a relatively high frequency in Caucasoids (0.18 ± 0.03). For the lipoprotein fragment length polymorphisms examined there was no reason to suspect that the difference in allele frequencies in affected and control groups was the result of foreign admixture. Indeed, APOA1 and LPL, the two markers whose allele frequencies

differed most markedly between Caucasoids and Micronesians did not show associations with hypercholesterolemia, hypertriglyceridemia or NIDDM. It is unlikely that population stratification can account for the APOC2/hypertriglyceridemia and APOD/NIDDM associations.

Even so, as the differences between lipoprotein allele frequencies between Micronesians and Caucasoids were not extreme, the lipoprotein markers may be poor indicators of the presence of foreign admixture. Though the Micronesian population is relatively homogeneous, the degree of foreign admixture present (12.1% of randomly selected, non-NIDDM adults had a demonstrably foreign HLA class I type, Serjeantson *et al.*, 1983) may be high enough to impinge on linkage analysis if there were different linkage disequilibrium relationships in chromosomal segments involved in predisposition to hypercholesterolemia, hypertriglyceridemia or NIDDM.

The power of RFLP data in linkage analysis may be improved if a map rather than single-markers is used. It is then possible to perform interval mapping, to test whether a putative locus lies between two markers whose locality has been previously determined, or to undertake a simultaneous search examining several candidate genes in concert (Lander and Botstein, 1986*b*). Even with the increased power obtained when using RFLP maps, it is felt that for heterogeneous disorders it would be preferable to obtain pedigrees with three affected individuals. Moreover, the use of RFLP maps in analysis does not overcome the

sensitivity of linkage analysis to errors of diagnosis (Lander and Botstein, 1986*b*).

5.6 CONCLUSIONS

1) Under a dominant model of inheritance the APOA1, APOC2, and APOD fragment length polymorphisms were excluded from linkage to hypercholesterolemia, hypertriglyceridemia and non-insulin dependent diabetes among Micronesians.

2) Under the same model the APOA2 fragment length polymorphism was excluded from linkage to hypertriglyceridemia, and the APOC1, and LPL RFLPs from linkage to hypercholesterolemia among Micronesians.

3) Under the same model linkage of the APOE fragment length polymorphism was excluded from linkage to hypercholesterolemia and non-insulin dependent diabetes among Micronesians.

4) Linkage of the APOD RFLP to hypercholesterolemia was excluded under a recessive model.

5) Linkage of the APOC2 RFLP to hypertriglyceridemia was excluded under a recessive model.

SECTION 6

GENERAL DISCUSSION

This thesis has examined the effect of genetic variation at eight lipoprotein loci on non-insulin dependent diabetes mellitus, obesity, central adiposity, hypercholesterolemia and hypertriglyceridemia. The *Taq*I site at the APOC2 locus was more often present in individuals with high plasma triglyceride levels and central adiposity. The APOD *Taq*I site was more often absent in individuals with NIDDM, high plasma cholesterol levels and central adiposity. The frequency of the APOE*2 allele was increased, with marginal significance, in NIDDM patients, but not in individuals with hypercholesterolemia or hypertriglyceridemia. However, linkage analysis has excluded linkage of the APOC2 and APOD loci with NIDDM, hypercholesterolemia and hypertriglyceridemia, modelled under dominant inheritance. The APOC2 and APOD loci where also excluded from linkage with hypertriglyceridemia and hypercholesterolemia respectively, when modelled under a model of recessive inheritance. In addition, the APOE locus has been excluded from linkage with NIDDM and hypercholesterolemia, under a model of dominant inheritance.

The apparent anomaly between a strong population association for APOD alleles and NIDDM and the exclusion of close linkage between APOD and a putative NIDDM locus under a dominant model of inheritance could indicate that the model for inheritance is excluded rather than linkage. Segregation analysis of NIDDM in Micronesians shows that a codominant model fits segregation patterns equally well as a dominant model (Serjeantson and Zimmet, 1989), suggesting a possible gene dosage effect. A gene dosage effect has

also been suggested for NIDDM by O'Rahilly *et al*. (1987). A gene dosage effect upon the expression of the disorder is modelled as for a dominant mode of inheritance, but with 50% penetrance in heterozygotes. The low penetrance value makes analyses, and lod scores, relatively uninformative.

Apparent anomalies between population and linkage analyses could also arise if an allele at the locus of interest was necessary, but not sufficient, for disease without some contribution from an additional unlinked gene. Then family members could share the susceptibility allele of interest but have different disease phenotypes. If this was the case those families monomorphic at the second locus could have high lod scores at the primary locus, while other families would not. This scheme is not inconsistent with the broad range of lod scores observed in families in this study for APOD and NIDDM, and for APOC2 and hypertriglyceridemia.

If this is the case, other loci of interest may need to be re-examined in the Micronesian population. Genes affecting insulin activity are obvious candidates, given the body of knowledge accumulating concerning their involvement in hyperinsulinemia and insulin-resistance. Hyperinsulinemia is associated with impaired glucose tolerance in many ethnic groups (Reaven and Miller, 1968; Sicree *et al*., 1987; Lillioja *et al*., 1988; Saad *et al*., 1988) and progression from a state of impaired glucose tolerance to diabetes is correlated with a decline in insulin response in Japanese (Kadowaki *et al*., 1984), Nauruans (Sicree *et al*., 1987) and Pima Indians (Saad *et al*., 1988).

Hyperinsulinemia has been shown to be associated with insulin resistance in the Pima Indians (Lillioja *et al.*, 1987) and both of these metabolic alterations are predictive of the development of diabetes (Knowler *et al.*, 1989). Hyperinsulinemia is also predictive of progression to diabetes among Nauruans with normal glucose tolerance (Sicree *et al.*, 1987).

Insulin resistance is a major causal factor for hyperglycaemia in NIDDM (Reaven *et al.*, 1976; Taylor *et al.*, 1990) and molecular defects have been identified in individuals with insulin resistance (Bell, 1991). These defects occur either in the insulin receptor gene, or in genes of the glucose transporter family. Seventeen mutations in patients with extreme insulin resistance have been identified (Bell, 1991), including missense and nonsense mutations (Kadowaki *et al.*, 1990) and deletions (Taira *et al.*, 1989). The mutations result either in the complete absence of the receptor molecule, or in defects in insulin receptor function, ranging from defective insulin binding and internalization of the insulin/insulin receptor complex to inefficient recycling of the receptor and impairment of tyrosine kinase activity. It has been suggested that if heterozygotes for these mutations are at risk for NIDDM, the gene may account for as much as 5-10% of the occurrence of NIDDM (Taylor *et al.*, 1990; Bell, 1991).

The family of glucose transporter proteins are responsible for the transport of glucose across the plasma membrane.

Facilitative glucose transporters are membrane proteins, present on the surface of all cells, which accelerate the passive and energy independent transport of glucose across the lipid bilayer (Bell *et al.*, 1990). Insulin has been shown to stimulate glucose transport by increasing the number of plasma membrane associated glucose transporters (Cushman and Wardzala, 1980; Suzuki and Kono, 1980) and so cause an increase in the rate of glucose uptake. However, in insulin resistant states there are fewer glucose transporters in both adipocytes and muscle cells (Simpson and Cushman, 1986).

Five functional glucose transporter genes and a glucose transporter pseudogene-like sequence have been identified (Bell *et al.*, 1990; Kayano *et al.*, 1990). Recent studies have indicated that decreased expression of the GLUT2 and GLUT4 genes may be responsible for insulin resistance in adipocytes and reduced insulin secretion in β-cells (Berger *et al.*, 1989; Sivitz *et al.*, 1989).

Molecular defects in the insulin receptor gene and altered expression of glucose transporters do not account for all cases of insulin resistance (Bell, 1991). However, progress has been made in identifying the defects and alterations listed above, and if, as previously suggested, insulin receptor gene defects contribute to insulin resistance in a subgroup of NIDDM patients, molecular screening techniques may be used to screen for further defects. The application of the polymerase chain reaction (Saiki *et al.*, 1985) and single stranded conformation and denaturing gradient

electrophoresis (Orita *et al.*, 1989; Sheffield *et al.*, 1989) greatly enhance the ability to detect molecular alterations and make more extensive molecular screening of diabetic individuals and at risk family members feasible.

Insulin resistance is also common in obese individuals and is often the result of a reduced number of glucose transporter molecules (Garvey *et al.*, 1988). There is a 36% decrease in GLUT4 mRNA in the subcutaneous adipocytes of obese compared to lean individuals. The amount of GLUT4 mRNA shows an even larger decline (86%) in NIDDM patients when compared with weight-matched controls (Garvey *et al.*, 1991). The suppression of GLUT4 expression in obesity does not account for all of the suppression observed in NIDDM. However, GLUT4 suppression is a major mechanism in insulin resistance in both disorders, and has also been observed in individuals with impaired glucose tolerance. Consequently, GLUT4 suppression may represent an early marker for progression to diabetes (Garvey *et al.*, 1991).

The suppression of GLUT4 expression is an interactional effect and does not appear related to any defects in the gene itself. A preliminary study of a GLUT4 *Kpn*I RFLP in Micronesians has not revealed any association of the locus with NIDDM (pers. comm. Prof. S.W. Serjeantson). Nor has the GLUT4 gene been implicated in the pathogenesis of NIDDM in the North American population (Bell *et al.*, 1990).

Although obesity is associated with an increased incidence of NIDDM (Kawate *et al.*, 1979; Stanhope and Prior, 1980 Mather and Keen, 1985), recent studies have indicated that

the distribution of body fat is an important factor in determining the nature of this association (Kissebah and Peiris, 1989). Diabetics have been shown to have more subscapular subcutaneous fat and a general excess of fat on the torso when compared with non-diabetics (Feldman *et al.*, 1969; Joos *et al.*, 1984).

The degree of central adiposity has also been shown to be related to insulin resistance. In a study of NIDDM in Japanese-American men both high fasting C-peptide levels, assumed to be present as the result of a compensation response to insulin resistance, and degree of central adiposity where found to be predictive of progression to diabetes. Moreover, the level of C-peptide was correlated with intra-abdominal fat area as determined by computer tomography (Bergstrom *et al.*, 1990), indicating that central adiposity may be associated with insulin resistance.

Both waist-to-hip and waist-to-thigh ratios have been used as an index of the degree of central adiposity and both have been shown to be associated with diabetes (Hartz *et al.*, 1984). Moreover, waist-to-hip ratio is predictive of both diabetes and cardiovascular disease (Ohlson *et al.*, 1985; Lapidus *et al.*, 1984), with the prediction of cardiovascular disease being independent of the overall level of obesity. The frequency of peripheral vascular disease, coronary heart disease and hypertension has been shown to be higher in diabetics with abdominal fat masses (van Gaal *et al.*, 1988). An increase in WHR has been found

to be associated with an increase in fasting plasma glucose and insulin levels, and with higher glucose and insulin responses to an oral glucose challenge. The hyperinsulinemia and peripheral resistance to glucose action found in subjects with abdominal obesity may lead to a decrease in activity of adipose lipoprotein lipase (Kissebah and Peiris, 1989).

The relationship of body fat distribution and diabetes has been observed in both men and women, over a range of ethnic groups and is also predictive of hyperlipoproteinemia (Kissebah and Peiris, 1989). In a study comparing waist-to-hip ratio with percentage of body fat and body mass index as correlates of plasma lipids, waist-to-hip ratio was positively associated with triglyceride, cholesterol, LDL-cholesterol and VLDL-cholesterol, and negatively associated with HDL-cholesterol. Moreover, after adjusting for the degree of central adiposity, overall obesity was not significantly associated with any of these variables (Terry *et al.*, 1989).

The WHR correlates highly with intra-abdominal visceral mass and it is suggested that the correlative power of WHR with diabetes, hypertension, cardiovascular disease and hyperlipoproteinemia is the result of its ability to predict the volume of abdominal visceral fat. Although central adiposity has been shown to be predictive of diabetes and hyperlipoproteinemia it is not suggested that the increased abdominal visceral fat mass is the primary causative agent in these metabolic alterations. Central

adiposity may be a visible early marker for diabetes and hyperlipoproteinemia at the same time as acting as an "environmental" stressor which may impact upon an underlying genetic predisposition towards poorly regulated hepatic glucose production with consequent decreased insulin secretion (Kissebah and Peiris, 1989).

Results presented in this thesis are in accordance with the general scheme of alterations in glucose and lipoprotein metabolism with increasing levels of central adiposity. The APOC2 locus was significantly associated with plasma triglyceride levels and the APOD locus with diabetic status. Both plasma triglyceride levels and diabetic status are associated with WHR. Consequently both the APOC2 and APOD loci were associated with WHR.

The limitations of single marker linkage analysis have been discussed (Section 5.5.2). Linkage analysis using genetic maps or highly variable markers will be required in future studies if the problems of detecting linkage with a single locus, against a background of genetic or polygenic heterogeneity are to be overcome. It is also clear that the application of a wider range of molecular techniques and the further clarification of the underlying biochemical mechanisms and inter-relationships of NIDDM, obesity and hyperlipoproteinemia are necessary if the primary genetic defects for these disorders are to be revealed.

REFERENCES

Abrams, J.J., H. Ginsberg, and S.M. Grundy. 1982. Metabolism of cholesterol and triglycerides in non-ketotic diabetes mellitus. *Diabetes* **31**:903-910.

Aburatani, H., T. Murase, F. Takaku, H. Itoh, A. Matsumoto, and H. Itakura. 1987. Apolipoprotein B-gene polymorphism and myocardial infarct. *N. Engl. J. Med.* **317**:52-55.

Alaupovic, P. 1971. Conceptual development of the classification systems of plasma lipoproteins. In *Protides of the Biological Fluids*. H. Peeters, editor. Pergamon Press, Oxford. pp. 9-20.

Albers, J.J., M.C. Cheung, S.L. Ewens, and J.H. Tollefson. 1981. Characterization and immunoassay of apolipoprotein D. *Atherosclerosis* **39**:395-409.

Anderson, R.A., T.J. Benda, R.B. Wallace, S.L. Eliason, J. Lee, and T.L. Burns. 1986. Prevalence and associations of apolipoprotein A-I linked DNA polymorphisms: results from a population study. *Genet. Epidemiol.* **3**:385-397.

Antonarakis, S.E., P. Oettgen, A. Chakravati, L. Feisse, and S.K. Karathanasis. 1986. DNA polymorphism analysis of the human apolipoprotein AI-CIII-AIV gene cluster. *Am. J. Hum. Genet.* **39**:Suppl. pp. 549.(Abstract).

Antonarakis, S.E., P. Oettgen, A. Chakravarti, S.L. Halloran, R.R. Hudson, L. Feisee, and S.K. Karathanasis. 1988. DNA polymorphism haplotypes of the human apolipoprotein APOA1-APOC3-APOA4 gene cluster. *Hum. Genet.* **80**:265-273.

Asakawa, J.N., N. Takahashi, B.B Rosenblum, and J.V. Neel. 1985. Two dimensional gel studies of genetic variation in the plasma proteins of Amerindians and Japanese. *Hum. Genet.* **50**:222-230.

Assmann, G. 1985. The function of human serum apolipoproteins: clues from inborn errors of lipoprotein metabolism. In *Human Apolipoprotein Mutants. Impact on Atherosclerosis and Longevity*. NATO advanced study institute series: Series A, Life sciences. Vol. 112. C.R. Sirtoli, A.V. Nichols, and G. Franceschini, editors. Plenum Press, New York. pp. 1-16.

Assmann, G., and H. Schulte. 1988. The Prospective Cardiovascular Munster (PROCAM) study: prevalence of hyperlipidemia in persons with hypertension and/or diabetes mellitus and the relationship to coronary heart disease. *Am. Heart J.* **116**:1713-1724.

Assmann, G., G. Schmitz, H.J. Menzel, and H. Schulte. 1984. Apolipoprotein E polymorphism and hyperlipidemia. *Clin. Chem.* **30**:641-643.

Assmann, G., A. von Eckardstein, H. Funke, and S. Rust. 1990. Characterization of apolipoprotein mutants. In *Atherosclerosis Reviews*. Vol. 20. A.M. Gotto Jr., and R. Paoletti, editors. Raven Press, New York. pp. 35-45.

Balkau, B., H. King, P. Zimmet, and L.R. Raper. 1985. Factors associated with the development of diabetes in the Micronesian population of Nauru. *Am. J. Epidemiol.* **122**:594-605.

Barker, W.C., and M.O. Dayhoff. 1977. Evolution of lipoproteins deduced from protein sequence data. *Comp. Biochem. Physiol. B.* **57**:309-315.

Barnett, A.H., C. Eff, R.D.G. Leslie, and D.A. Pyke. 1981. Diabetes in identical twins: a study of 200 pairs. *Diabetologia* **20**:87-93.

Bell, G.I. 1991. Molecular genetics in diabetes mellitus. *Diabetes* **40**:413-422.

Bell, G.I., T. Kayano, J.B. Buse, C.F. Burant, J. Takeda, D. Lin, H. Fukumoto, and S. Seino. 1990. Molecular biology of the mammalian glucose transporters. *Diabetes Care* **13**:198-208.

Bell, G.I., K.S. Xiang, M.V. Newman, S.H. Wu, L.G. Wright, S.S. Fajans, R.S. Spielman, and N.J. Cox. 1991. Gene for non-insulin-dependent diabetes mellitus (maturity-onset diabetes of the young subtype) is linked to DNA polymorphism on human chromosome 20q. *Proc. Natl. Acad. Sci. U. S. A.* **88**:1484-1488.

Berg, K. 1990. Risk factor variability and coronary heart disease. *Acta Genet. Med. Gemellol. Roma.* **39**:15-24.

Berg, K., L.M. Powell, S.C. Wallis, R. Pease, T.J. Knott, and J. Scott. 1986. Genetic linkage between the antigenic group (Ag) variation and the apolipoprotein B gene: assignment of the Ag locus. *Proc. Natl. Acad. Sci. U. S. A.* **83**:7367-7370.

Berger, J., C. Biswas, P.P. Vicario, H.V. Strout, R. Saperstein, and P.F. Pilch. 1989. Decreased expression of the insulin-responsive glucose transporter in diabetes and fasting. *Nature* **340**:70-72.

Bergstrom, R.W., L.L. Newell-Morris, D.L. Leonetti, W.P. Shuman, P.W. Wahl, and W.Y. Fujimoto. 1990. Association of elevated fasting C-peptide level and increased intra-abdominal fat distribution with development of NIDDM in Japanese-American men. *Diabetes* **39**:104-111.

Betteridge, D.J. 1986. Lipoprotein metabolism. In *Recent Advances in Diabetes 2*. M. Nattrass, editor. Churchill Livingstone, New York, pp. 91-107.

Betteridge, D.J. 1989. Diabetes, lipoprotein metabolism and atherosclerosis. *Br. Med. Bull.* **45**:285-311.

Bisgaier, C.L., E.S. Lee, and R.M. Glickman. 1987. A method to screen apolipoprotein polymorphisms in whole plasma: description of apolipoprotein A-IV in dyslipidemias and a reassessment of apolipoprotein A-I in Tangier disease. *Biochim. Biophys. Acta* **918**:242-249.

Black, S.C., S. Hewett, Y. Kotubi, R.V. Brunt, and J.P. Reckless. 1990. Isoform patterns of apolipoprotein E in diabetes mellitus. *Diabetic Med.* **7**:532-539.

Blackhart, B.D., E.M. Ludwig, V.R. Pierotti, L. Caiati, M.A. Onasch, S.C. Wallis, L. Powell, R. Pease, T.J. Knott, M.L. Chu, R.W. Mahley, J. Scott, B.J. McCarthy and B. Levy-Wilson. 1986. Structure of the human apolipoprotein B gene. *J. Biol. Chem.* **261**:15364-15367.

Bodmer, W.F., C.J. Bailey, J. Bodmer, H.J.R. Bussey, A. Ellis, P. Gorman, F.C. Lucibello, V.A. Murday, S.H. Rider, P. Scambler, D. Sheer, E. Solomon, and N.K. Spurr. 1987. Localization of the gene for familial adenomatous polyposis in chromosome 5. *Nature* **328**:614-616.

Boerwinkle, E., and G. Utermann. 1988. Simultaneous effects of the apolipoprotein E polymorphism on apolipoprotein E, apolipoprotein B, and cholesterol metabolism. *Am. J. Hum. Genet.* **42**:104-112.

Boerwinkle, E., S. Visvikis, D. Welsh, J. Steinmetz, S.M. Hanash, and C.F. Sing. 1987. The use of measured genotype information in the analysis of quantitative phenotypes in man. II. The role of the apolipoprotein E polymorphism in determining levels, variability, and covariability of cholesterol, betalipoprotein, and triglycerides in a sample of unrelated individuals. *Am. J. Med. Genet.* **27**:567-582.

Botstein, D., R. White, M. Skolnick, and R. Davis. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* **32**:314-331.

Breslow, J.L. 1985. Human apolipoprotein molecular biology and genetic variation. *Annu. Rev. Biochem.* **54**:699-727.

Breslow, J.L. 1988. Apolipoprotein genetic variation and human disease. *Physiol. Rev.* **68**:85-132.

Breslow, J.L., V.I. Zannis, T.R. SanGiacomo, J.L. Third, T. Tracy, and C.J. Glueck. 1982*a*. Studies of familial type III hyperlipoproteinemia using as a genetic marker the apoE phenotype E2/2. *J. Lipid Res.* **23**:1224-1235.

Breslow, J.L., D. Ross, J. McPherson, H. Williams, D. Kurnit, A.L. Nussbaum, S.K. Karathanasis, and V.I. Zannis. 1982*b*. Isolation and characterization of cDNA clones for human apolipoprotein A-I. *Proc. Natl. Acad. Sci. U. S. A.* **79**:6861-6865.

Brewer, H.B.Jr., S.E. Lux, R. Ronan, and K.M. John. 1972. Amino acid sequence of human apoLp-Gln-II (apoA-II), an apolipoprotein isolated from the high density lipoprotein complex. *Proc. Natl. Acad. Sci. U. S. A.* **69**:1304-1308.

Brewer, H.B.Jr., R. Shulman, P. Herbert, R. Ronan, and K. Wehrly. 1974. The complete amino sequence of alanine apolipoprotein (apoC-III), an apolipoprotein from human plasma very low density lipoproteins. *J. Biol. Chem.* **249**:4975-4984.

Brink, P.A., L.T. Steyn, A.J. Bester, and K. Steyn. 1986. Linkage disequilibrium between a marker on the low-density lipoprotein receptor and high cholesterol levels. *S. Afr. Med. J.* **70**:80-82.

Brink, P.A., L.T. Brink, M. Torrington, and A.J. Bester. 1990. Linkage study of the low-density lipoprotein-receptor gene and cholesterol levels in an Afrikaner family. Quantitative genetics and identification of a minor founder effect. *S. Afr. Med. J.* **77**:292-296.

Brosseau, J.D., R.C. Eelkema, A.C. Crawford, and T.A. Abe. 1979. Diabetes among the three affiliated tribes: correlation with degree of Indian inheritance. *Am. J. Public Health* **69**:1277-1278.

Brown, M.S., and J.L. Goldstein. 1988. The LDL receptor concepts: clinical and therapeutic implications. In *Atherosclerosis Reviews*. Vol. 18. A.M. Gotto Jr., and R. Paoletti, editors. Raven Press, New York. pp. 85-93.

Brunzell, J.D., A. Chait, and E.L. Bierman. 1985. Plasma lipoproteins in human diabetes mellitus. In *The Diabetes Annual/1*. K.G.M.M. Alberti, and L.P. Krall, editors. Elsevier Science Publishers, Amsterdam. pp. 463-479.

Buraczynska, M., J. Hanzlik, and M. Grzywa. 1985. Apolipoprotein A-I gene polymorphism and susceptibility of non-insulin-dependent diabetes mellitus. *Am. J. Hum. Genet.* **37**:1129-1137.

Buraczynska, M., J. Hanzlik, and M. Grzywa. 1986. ApoA-I related DNA polymorphism in humans with coronary heart disease. *Hum. Genet.* **74**:165-167.

Chakraborty, R., R.E. Ferrell, M.P. Stern, S.M. Haffner, H.P. Hazuda, and M. Rosenthal. 1986. Relationship of prevalence of non-insulin dependent diabetes mellitus with Amerindian admixture in the Mexican Americans of San Antonio, Texas. *Genet. Epidemiol.* **3**:435-454.

Chamberlain, J.C., J.A. Thorn, K. Oka, D.J. Galton, and J. Stocks. 1989. DNA polymorphisms at the lipoprotein lipase gene: associations in normal and hypertriglyceridaemic subjects. *Atherosclerosis* **79**:85-91.

Chotai, J. 1984. On the lod score method in linkage analysis. *Ann. Hum. Genet.* **48**:359-378.

Cole, S.A., E.J. Szathmary, and R.E. Ferrell. 1989. Gene and gene-product variation in the apolipoprotein A-I/C-III/A-IV cluster in the Dogrib Indians of the Northwest Territories. *Am. J. Hum. Genet.* **44**:835-843.

Coleman, R.T., P.A. Gonzalez, H. Funke, G. Assmann, B. Levy Wilson, and P.M. Frossard. 1986. Polymorphisms in the apolipoprotein AI-CIII gene complex. *Mol. Biol. Med.* **3**:213-228.

Collins, A., and N.E. Morton. 1991. Significance of maximal lods. *Ann. Hum. Genet.* **55**:39-41.

Collins, V.R., G.K. Dowse, C.F. Finch, P.Z. Zimmet, and A.W. Linnane. 1989. Prevalence and risk factors for micro- and macroalbuminuria in diabetic subjects and entire population of Nauru. *Diabetes* **38**:1602-1610.

Cooper, D.N., and J.F. Clayton. 1988. DNA polymorphism and the study of disease associations. *Hum. Genet.* **78**:299-312.

Crouse, J., and D. Amorese. 1987. Ethanol precipitation: ammonium acetate as an alternative to sodium acetate. *Focus* **9**:3-5.

Cumming, A.M., and F.W. Robertson. 1984. Polymorphism at the apoprotein-E locus in relation to risk of coronary disease. *Clin. Genet.* **25**:310-313.

Curtiss, L.K., and J.L. Witztum. 1983. A novel method for generating region-specific monoclonal antibodies to modified proteins. Application to the identification of human glycosylated low density lipoproteins. *J. Clin. Invest.* **72**:1427-1438.

Curtiss, L.K., and J.L. Witztum. 1985. Plasma apolipoproteins A-I, A-II, B, C-I and E are glycosylated in hyperglycaemic diabetic subjects. *Diabetes* **34**:452-461.

Cushman, S.W., and L.J. Wardzala. 1980. Potential mechanism of insulin action on glucose transport in the isolated rat adipose cell: apparent translocation of intracellular transport systems to the plasma membrane. *J. Biol. Chem.* **225**:4758-4762.

Daga, A., M. Fabbi, T. Mattioni, S. Bertolini, and G. Corte. 1988. *Pvu*II polymorphism of low density lipoprotein receptor gene and familial hypercholesterolemia. Study of Italians. *Arteriosclerosis* **8**:845-850.

Dagert, M., and S.D. Ehrlich. 1977. Prolonged incubation in calcium chloride improves the competence of *Escherichia coli* cells. *Gene* **6**:23-28.

Das, H.K., J. McPherson, G.A. Bruns, S.K. Karathanasis, and J.L. Breslow. 1985. Isolation, characterization, and mapping to chromosome 19 of the human apolipoprotein E gene. *J. Biol. Chem.* **260**:6240-6247.

Davignon, J., R.E. Gregg, and C.F. Sing. 1988. Apolipoprotein E polymorphism and atherosclerosis. *Arteriosclerosis* **8**:1-21.

Deeb, S.S., and R. Peng. 1989. Structure of the human lipoprotein lipase gene. *Biochemistry* **28**:4131-4135.

Deeb, S., A. Failor, B.G. Brown, J.D. Brunzell, J.J. Albers, and A.G. Motulsky. 1986. Molecular genetics of apolipoproteins and coronary heart disease. *Cold Spring Harb. Symp. Quant. Biol.* **51** (Pt 1):403-409.

de Loof, H., M. Rosseneu, C.Y. Yang, W.H. Li, A.M. Gotto Jr., and L. Chan. 1987. Human apolipoprotein B: analysis of internal repeats and homology with other apolipoproteins. *J. Lipid Res.* **28**:1455-1465.

Devlin, R.H., S. Deeb, J. Brunzell, and M.R. Hayden. 1990. Partial gene duplication involving exon-*Alu* interchange results in lipoprotein lipase deficiency. *Am. J. Hum. Genet.* **46**:112-119.

Dowse, G.K., P.Z. Zimmet, C.F. Finch, V.R. Collins. 1991. Decline in the incidence of epidemic glucose intolerance in Nauruans: implications for the "Thrifty genotype". *Am. J. Epidemiol.* **133**:1093-1104.

Drayna, D., C. Fielding, J. McLean, B. Baer, G. Castro, E. Chen, L. Comstock, W. Henzel, W. Kohr, L. Rhee, K. Wion and R. Lawn. 1986. Cloning and expression of human apolipoprotein D cDNA. *J. Biol. Chem.* **261**:16535-16539.

Drayna, D.T., J.W. McLean, K.L. Wion, J.M. Trent, H.A. Drabkin, and R.M. Lawn. 1987*a*. Human apolipoprotein D gene: gene sequence, chromosome localization, and homology to the α_{2u}-gobulin superfamily. *DNA* **6**:199-204.

Drayna, D., J.D. Scott, and R. Lawn. 1987*b*. Multiple RFLPs at the human apolipoprotein D (APOD) locus. *Nucleic Acids Res*. **15**:9617.

Easteal, S., A.H. Voohoeve, P. Zimmet and S.W. Serjeantson. 1989. RFLP-defined HLA-DP polymorphism in four ethnic groups. *Hum. Immunol*. **25**:169-179.

Ehnholm, C., M. Lukka, T. Kuusi, E. Nikkila, and G. Utermann. 1986. Apolipoprotein E polymorphism in the Finnish population: gene frequencies and relation to lipoprotein concentrations. *J. Lipid Res*. **27**:227-235.

Elbein, S.C., L. Corsetti, A. Ullrich, and M.A. Permutt. 1986. Multiple restriction fragment length polymorphisms at the insulin receptor locus: a highly informative marker for linkage analysis. *Proc. Natl. Acad. Sci. U. S. A.* **83**:5223-5227.

Elshourbagy, N.A., D.W. Walker, M.S. Boguski, J.I. Gordon, and J.M. Taylor. 1986. The nucleotide and derived amino acid sequence of human apolipoprotein A-IV mRNA and the close linkage of its gene to the genes of apolipoprotein A-I and C-III. *J. Biol. Chem*. **162**: 1990-2002.

Elshourbagy, N.A., D.W. Walker, Y.K. Paik, M.S. Boguski, M. Freeman, J.I. Gordon, and J.M. Taylor. 1987. Structure and expression of the human apolipoprotein A-IV gene. *J. Biol. Chem*. **262**:7973-7981.

Elston, R.C., K.K. Namboodiri, H.V. Nino, and W.S. Pollitzer. 1974. Studies on blood and urine glucose in Seminole Indians: indications for segregation of a major gene. *J. Hum. Genet*. **26**:13-34.

Emi, M., L.L. Wu, M.A. Robertson, R.L. Myers, R.A. Hegele, R.R. Williams, R. White, and J.M. Lalouel. 1988. Genotyping and sequence analysis of apolipoprotein E isoforms. *Genomics* **3**:373-379.

Emi, M., A. Hata, M. Robertson, P.H. Iverius, R. Hegele, and J.M. Lalouel. 1990*a*. Lipoprotein lipase deficiency resulting from a nonsense mutation in exon 3 of the lipoprotein lipase gene. *Am. J. Hum. Genet*. **47**:107-111.

Emi, M., D.E. Wilson, P.H. Iverius, L. Wu, A. Hata, R. Hegele, R.R. Williams, and J.M. Lalouel. 1990*b*. Missense mutation (Gly-Glu188) of human lipoprotein lipase imparting functional deficiency. *J. Biol. Chem*. **265**:5910-5916.

Eto, M., K. Watanabe, and K. Ishii. 1986*a*. A racial difference in apolipoprotein E allele frequencies between the Japanese and Caucasian populations. *Clin. Genet*. **30**:442-427.

Eto, M., K. Watanabe, Y. Iwashima, A. Morikawa, E. Oshima, M. Sekiguchi, and K. Ishii. 1986*b*. Apolipoprotein E polymorphism and hyperlipemia in type II diabetics. *Diabetes* **35**:1374-1382.

Eto, M., K. Watanabe, Y. Iwashima, A. Morikawa, N. Chonan, E. Oshima, M. Sekiguchi, and K. Ishii. 1987. Increased frequency of apolipoprotein epsilon 4 allele in type II diabetes with hypercholesterolemia. *Diabetes* **36**:1301-1306.

Eto, M., K. Watanabe, and K. Ishii. 1989. Apolipoprotein E polymorphism and hyperlipoproteinemia in obesity. *Int. J. Obes.* **13**:433-440.

Falus, A., and L. Romics. 1988. New approaches in medical molecular genetics: presymptomatic diagnosis in familial hypercholesterolemia by LDL receptor cDNA probe. *Clin. Chem. Enzym. Comms.* **1**:35-38.

Feinberg, A.P., and B. Vogelstein. 1983. A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **132**:6-13.

Feinberg, A.P., and B. Vogelstein. 1984. "A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity". Addendum. *Anal. Biochem.* **137**:266-267.

Feldman, R., A.J. Sender, and A.B. Seigelaub. 1969. Difference in diabetic and non-diabetic fat distribution patterns by skinfold measurement. *Diabetes* **18**:478-486.

Ferns, G.A., and D.J. Galton. 1986*a*. Haplotypes of the human apoprotein AI-CIII-AIV gene cluster in coronary atherosclerosis. *Hum. Genet.* **73**:245-249.

Ferns, G.A., and D.J. Galton. 1986*b*. Frequency of *Xba*I polymorphism in myocardial infarct survivors [letter]. *Lancet* **ii**:572.

Ferns, G.A., J. Stocks, C. Ritchie, and D.J. Galton. 1985. Genetic polymorphisms of apolipoprotein C-III and insulin in survivors of myocardial infarction. *Lancet* **ii**:300-303.

Ferns, G.A., J. Stocks, and D.J. Galton. 1986*a*. C-III DNA restriction fragment length polymorphism and myocardial infarction [letter]. *Lancet* **i**:94.

Ferns, G.A., C.S. Shelley, J. Stocks, A. Rees, H. Paul, F. Baralle, and D.J. Galton. 1986*b*. A DNA polymorphism of the apoprotein AII gene in hypertriglyceridaemia. *Hum. Genet.* **74**:302-306.

Fielding, P.E., and C.J. Fielding. 1980. A cholesterol transfer complex in human plasma. *Proc. Natl. Acad. Sci. U. S. A.* **77**:3327-3330.

Fisher, K.L., G.A. FitzGerald, and R.M. Lawn. 1987. Two polymorphisms in the human lipoprotein lipase (LPL) gene. *Nucleic Acids Res.* **15**:7657.

Flint, J., A.J. Boyce, J.J. Martinson, and J.B. Clegg. 1989. Population bottlenecks in Polynesia revealed by minisatellites. *Hum. Genet.* **83**:257-263.

Fojo, S.S., A.F. Stalenhoef, K. Marr, R.E. Gregg, R.S. Ross, and H.B. Brewer Jr. 1988*a*. A deletion mutation in the ApoC-II gene (ApoC-II Nijmegen) of a patient with a deficiency of apolipoprotein C-II. *J. Biol. Chem.* **263**:17913-17916.

Fojo, S.S., U. Beisiegel, U. Beil, K. Higuchi, M. Bojanovski, R.E. Gregg, H. Greten, and H.B. Brewer Jr. 1988*b*. Donor splice site mutation in the apolipoprotein (Apo) C-II gene (Apo C-II Hamburg) of a patient with Apo C-II deficiency. *J. Clin. Invest.* **82**:1489-1494.

Freeman, D.J., C.J. Packard, J. Shepherd, and D. Gaffney. 1990. Polymorphisms in the gene coding for cholesteryl ester transfer protein are related to plasma high-density lipoprotein cholesterol and transfer protein activity. *Clin. Sci.* **79**:575-581.

Frezal, J., and H.P. Klinger. 1987. Human gene mapping 9. *Cytogenet. Cell Genet.* Vol. 46 (No. 1-4).

Frossard, P.M., H. Funke, R.T. Coleman, and G. Assmann. 1986. Genetic markers for coronary atherosclerosis in the human apolipoprotein AI-CIII-AIV gene complex. *Am. J. Hum. Genet.* **39**:Suppl. pp. 589.(Abstract).

Frossard, P.M., R.T. Coleman, M.J. Malloy, J.P. Kane, B. Levy Wilson, and V.A. Appleby. 1987. Human apolipoprotein CI (apoCI) gene locus: *Dra*I dimorphic site. *Nucleic Acids Res.* **15**:1884.

Fukumoto, H., S. Seino, H. Imura, Y. Seino, R.L. Eddy, Y. Fukushima, M.C. Byers, T.B. Shows, and G.I. Bell. 1988. Sequence, tissue distribution, and chromosomal localization of mRNA encodng a human glucose transporter-like protein. *Proc. Natl. Acad. Sci. U. S. A.* **85**:5434-5438.

Fumeron, F., D. Rigaud, M.C. Bertiere, S. Bardon, C. Dely, and M. Apfelbaum. 1988. Association of apolipoprotein epsilon 4 allele with hypertriglyceridemia in obesity. *Clin. Genet.* **34**:258-264.

Funke, H., S. Rust, and G. Assmann. 1986. Detection of apolipoprotein E variants by an oligonucleotide "melting" procedure. *Clin. Chem.* **32**:1285-1289.

Galton, D.J. 1985. The hyperlipidemias. In *Molecular Genetics of Common Metabolic Disease*. D.J. Galton, editor. Edward Arnold Publishers, London. pp. 83-99.

Galton, D.J. 1987. A genetic analysis of premature coronary artery disease. In *New Approaches to Genetic Diseases*. T. Sasazuki, editor. Academic Press, London. pp. 3-11.

Galton, D.J. 1990. DNA variation of the apolipoprotein loci and risk for coronary atherosclerosis. In *From Phenotypes to Genes in Common Disorders*. K. Berg, N. Retterstol, and S. Refsum, editors. Munksgaard International Publishers, Copenhagen. pp. 60-76.

Garfinkel, A.S., and Schotz, M.C. 1987. Lipoprotein lipase. In *Plasma Lipoproteins*. A.M. Gotto Jr., editor. Elsevier Sciene Publishers, New York. pp. 335-357.

Garvey, W.T., T.P. Heucksteadt, S. Matthaei, and J.M. Olefsky. 1988. The role of glucose transporters in the cellular insulin resistance of type II non-insulin dependent diabetes mellitus. *J. Clin. Invest.* **81**:1528-1536.

Garvey, W.T., L. Maianu, T.P. Heucksteadt, M. Birnbaum, and J.M. Molina. 1991. Pretranslational suppression of a glucose transporter protein causes insulin resistance in adipocytes from patients with non-insulin dependent diabetes mellitus and obesity. *J. Clin. Invest.* **87**:1072-1081.

Ghiselli, G., J. Schaefer, L.A. Zech, R.E. Gregg, and H.B. Brewer Jr. 1982. Increased prevalence of apolipoprotein E4 in type V hyperlipoproteinemia. *J. Clin. Invest.* **70**:474-477.

Gibbons, G.F. 1986. Hyperlipidaemia of diabetes. *Clin. Sci.* **71**:477-486.

Glomset, J.A. 1968. The plasma lecithin:cholesterol acyltransferase reaction. *J. Lipid Res.* **9**:155-167.

Goldstein, J.L., and M.S. Brown. 1983. Familial hypercholesterolemia. In *The Metabolic Basis of Inherited Disease*. J.B. Stanbury, J.B. Wyngarrden, D.S. Fredrickson, J.L. Golstein, and M.S. Brown, editors. McGraw-Hill, New York. pp. 672-712.

Goldstein, J.L., B.A. Kottke, and M.S. Brown. 1982. Biochemical genetics of LDL receptor mutations in familial hypercholesterolemia. *Prog. Clin. Biol. Res.* **103** (Pt B):161-176.

Grunebaum, L., J.P. Cazenave, G. Camerino, C. Kloepfer, J.L. Mandel, P. Tolstoshev, M. Jaye, H. de la Salle, and J.P. Lecocq. 1984. Carrier detection of Hemophilia B by using a restriction site polymorphism associated with the coagulation Factor IX gene. *J. Clin. Invest.* **73**:1491-1495.

Gueguen, R., S. Visvikis, J. Steinmetz, G. Siest, and E. Boerwinkle. 1989. An analysis of genotype effects and their interactions by using the apolipoprotein E polymorphism and longitudinal data. *Am. J. Hum. Genet.* **45**:793-802.

Gurr, M.I., and A.T. James. 1980. *Lipid Biochemistry: An Introduction*. Chapman and Hall, London. pp. 247.

Haldane, J.B.S., and C.A.B. Smith. 1947. A new estimate of the linkage between the genes for colour-blindness and haemophilia in man. *Ann. Eugen.* **14**:10-31.

Hamaguchi, H., N. Hattori, J. Satoh, M. Onuki, K. Yamakawa, H. Fujiwara, H. Amamiya, H. Nagaoka, Y. Yamanouchi, T. Okafuji, Y. Iwamura, and S. Tsuchiya. 1987. Apolipoprotein AI-CIII gene polymorphisms and myocardial infarction in a Japanese population. In *New Approaches to Genetic Diseases*. T. Sasazuki, editor. Academic Press, London. pp. 14-21.

Hanis, C.L., D. Hewett Emmett, T.C. Douglas, and W.J. Schull. 1991. Lipoprotein and apolipoprotein levels among Mexican-Americans in Starr County, Texas. *Arterioscler. Thromb.* **11**:123-129.

Hartz, A.J., D.C. Rupley, and A.A. Rimm. 1984. The association of girth measurements with disease in 32,856 women. *Am. J. Epidemiol.* **119**:71-80.

Hata, A., M. Emi, G. Luc, A. Basdevant, P. Gambert, P.H. Iverius, and J.M. Lalouel. 1990*a*. Compound heterozygote for lipoprotein lipase deficiency: Ser-Thr244 and transition in 3′ splice site of intron 2 (AG-AA) in the lipoprotein lipase gene. *Am. J. Hum. Genet.* **47**:721-726.

Hata, A., M. Robertson, M. Emi, and J.M. Lalouel. 1990*b*. Direct detection and automated sequencing of individual alleles after electrophoretic strand separation: identification of a common nonsense mutation in exon 9 of the human lipoprotein lipase gene. *Nucleic Acids Res.* **18**:5407-5411.

Hayden, M.R., H. Kirk, C. Clark, J. Frohlich, S. Rabkin, R. McLeod, and J. Hewitt. 1987. DNA polymorphisms in and around the Apo-A1-CIII genes and genetic hyperlipidemias. *Am. J. Hum. Genet.* **40**:421-430.

Healy, M.J.R. 1988. *GLIM: An Introduction*. Clarendon Press, Oxford. pp. 130.

Hegele, R.A., and J.L. Breslow. 1987. Apolipoprotein genetic variation in the assessment of atherosclerosis susceptibility. *Genet. Epidemiol.* **4**:163-184.

Hegele, R.A., L.S. Huang, P.N. Herbert, C.B. Blum, J.E. Buring, C.H. Hennekens, and J.L. Breslow. 1986. Apolipoprotein B-gene DNA polymorphisms associated with myocardial infarction. *N. Engl. J. Med.* **315**:1509-1515.

Hegele, R.A., C.H. Hennekens, and J.L. Breslow. 1989. Allele frequencies of apolipoprotein A-I and A-II gene locus DNA polymorphisms in Boston-based whites. *Hum. Hered.* **39**:174-178.

Hegele, R.A., R. Plaetke, and J.M. Lalouel. 1990. Linkage disequilibrium between DNA markers at the low-density lipoprotein receptor gene. *Genet. Epidemiol.* **7**:69-81.

Herbert, P.N., G. Assmann, A.M. Gotto Jr., and D.S. Fredrickson. 1983. Familial lipoprotein deficiency: Abetalipoproteinemia, hypolipoproteinemia and Tangiers Disease. In *The Metabolic Basis of Inherited Disease*. J.B. Stanbury, J.B. Wyngaarden, D.S. Fredrickson, J.L. Golstein, and M.S. Brown, editors. McGraw-Hill, New York. pp. 589-621.

Hertzberg, M.S., K.N.P. Mickelson, and R.J. Trent. 1988. α-globin gene haplotypes in Polynesians: their relationships to population groups and gene arrangements. *Am. J. Hum. Genet.* **43**:971-977.

Hiramatsu, K., E.L. Bierman, and A. Chait. 1985. Metabolism of low-density lipoprotein from patients with diabetic hypertriglyceridaemia by cultured skin fibroblasts. *Diabetes* **34**:8-14.

Hixson, J.E., and D.T. Vernier. 1990. Restriction isotyping of human apolipoprotein E by gene amplification and cleavage with *Hha*I. *J. Lipid Res.* **31**:545-548.

Hobbs, H.H., M.A. Lehrman, T. Yamamoto, and D.W. Russell. 1985. Polymorphism and evolution of *Alu* sequences in the human low density lipoprotein receptor gene [published erratum appears in Proc. Natl. Acad. Sci. U. S. A. 1986 Mar;83(6):1964]. *Proc. Natl. Acad. Sci. U. S. A.* **82**:7651-7655.

Hobbs, H.H., V. Esser, and D.W. Russell. 1987. *Ava*II polymorphism in the human LDL receptor gene. *Nucleic Acids Res.* **15**:379-379

Hodge, S.E., C.E. Anderson, K. Neiswanger, R.S. Sparkes, and D.L. Remoin. 1983. The search for hereogeneity in insulin-dependent diabetes mellitus (IDDM): linkage studies, two-locus models, and genetic heterogeneity. *Am. J. Hum. Genet.* **35**:1139-1153.

Howard, B.V. 1987. Lipoprotein metabolism in diabetes mellitus. *J. Lipid Res.* **28**:613-628.

Huang, L-S., S.C. Bock, S.I. Feinstein, and J.L. Breslow. 1985. Human apolipoprotein B cDNA clone isolation and demonstration that liver apolipoprotein B mRNA is 22 kilobases in length. *Proc. Natl. Acad. Sci. U. S. A.* **82**:6825-6829.

Humphries, S.E., N.I. Jowett, L. Williams, A. Rees, M. Vella, A. Kessling, O. Myklebost, A. Lydon, M. Seed, D.J. Galton, and R. Williamson. 1983. A DNA polymorphism adjacent to the human apolipoprotein CII gene. *Mol. Biol. Med.* **1**:463-471.

Humphries, S.E., L. Williams, O. Myklebost, A.F. Stalenhoef, P.N. Demacker, G. Baggio, G. Crepaldi, D.J. Galton, and R. Williamson. 1984. Familial apolipoprotein CII deficiency: a preliminary analysis of the gene defect in two independent families. *Hum. Genet.* **67**:151-155.

Humphries, S.E., A.M. Kessling, B. Horsthemke, J.A. Donald, M. Seed, N. Jowett, M. Holm, D.J. Galton, V. Wynn, and R. Williamson. 1985. A common DNA polymorphism of the low-density lipoprotein (LDL) receptor gene and its use in diagnosis. *Lancet* **i**:1003-1005.

Humphries, S.E., P.J. Talmud, and A.M. Kessling. 1987. Use of DNA polymorphisms of the apolipoprotein genes to study the role of genetic variation in the determination of serum lipid levels. *Ciba. Found. Symp.* **130**:128-149.

Humphries, S., D.A. Coviello, P. Masturzo, R. Balestreri, G. Orecchini, and S. Bertolini. 1991. Variation in the low density lipoprotein receptor gene is associated with differences in plasma low density lipoprotein levels in young and old normal individuals from Italy. *Arteriosclerosis* **11**:509-516.

Imari, Y., S. Koga, and H. Ibayashi. 1988. Phenotypes of apolipoprotein E and abnormalities in lipid metabolism in patients with non-insulin-dependent diabetes mellitus. *Metabolism* **37**:1134-1138.

Innerarity, T.L., and R.W. Mahley. 1978. Inhanced binding by cultured human fibroblasts of apo-E containing lipoproteins compared with low density lipoproteins. *Biochemistry* **17**:1440-1447.

Ish-Horowicz, D., and J.F. Burke. 1981. Rapid and efficient cosmid cloning. *Nucleic Acids Res.* **9**:2989-2998.

Jackson, C.L., G.A. Bruns, and J.L. Breslow. 1984. Isolation and sequence of a human apolipoprotein CII cDNA clone and its use to isolate and map to human chromosome 19 the gene for apolipoprotein CII. *Proc. Natl. Acad. Sci. U. S. A.* **81**:2945-2949.

Jahn, C.E., J.C. Osborne Jr., E.J. Schaefer, and H.B. Brewer Jr. 1983. Activation of enzymatic activity of hepatic lipase by apolipoprotein A-II. *Eur. J. Biochem.* **131**:25-29.

Janus, E.D., A.M. Nicoll, P.R. Turner, P. Magill, and B. Lewis. 1980. Kinetic basis of the primary hyperlipidemias of apolipoprotein-B turnover in genetically defined subjects. *Eur. J. Clin. Invest.* **10**:161-172.

Joos, S.K., W.H. Mueller, C.L. Hanis, and W.J. Schull. 1984. Diabetes alert study: weight history and upper body obesity in diabetic and non-diabetic Mexican American adults. *Ann. Hum. Biol.* **11**:167-171.

Jowett, N.I., A. Rees, L.G. Williams, J. Stocks, M.A. Vella, G.A. Hitman, J. Katz, and D.J. Galton. 1984. Insulin and apolipoprotein A-1/C-III gene polymorphisms relating to hypertriglyceridaemia and diabetes mellitus. *Diabetologia* **27**:180-183.

Kadowaki, T., Y. Miuake, R. Hagura, Y. Akanuma, H. Kajinuma, N. Kuzuya, and F. Takaku. 1984. Risk factors for worsening diabetes in subjects with impaired glucose tolerance. *Diabetologia* **26**:44-49.

Kadowaki, T., Kadowaki, H., M.M. Rechler, M. Serrano-Rios, J. Roth, P. Gorden, and S.I. Taylor. 1990. Five mutant alleles in the insulin receptor gene in patients with genetic forms of insulin resistance. *J. Clin. Invest.* **86**:256-264.

Kamboh, M.I., and R.E. Ferrell. 1987. Genetic studies of human apolipoproteins. I. Polymorphism of apolipoprotein A-IV. *Am. J. Hum. Genet.* **41**:119-127.

Kamboh, M.I., R.E. Ferrell, and B. Sepehrnia. 1988*a*. Genetic studies of human apolipoproteins. IV. Structural heterogeneity of apolipoprotein H (beta 2-glycoprotein I). *Am. J. Hum. Genet.* **42**:452-457.

Kamboh, M.I., R.E. Ferrell, and B. Kottke. 1988*b*. Genetic studies of human apolipoproteins. V. A novel rapid procedure to screen apolipoprotein E polymorphism. *J. Lipid Res.* **29**:1535-1543.

Kamboh, M.I., B. Sepehrnia, and R.E. Ferrell. 1989. Genetic studies of human apolipoproteins. VI. Common polymorphism of apolipoprotein E in blacks. *Dis. Markers* **7**:49-55.

Kamboh, M.I., K.K. Bhatia, and R.E. Ferrell. 1990. Genetic studies of human apolipoproteins. XII. Population genetics of apolipoprotein polymorphisms in Papua New Guinea. *Am. J. Hum. Biol.* **2**:117-123.

Kamboh, M.I., K.M. Weiss, and R.E. Ferrell. 1991. Genetic studies of human apolipoproteins. XVI: APOE polymorphism and cholesterol levels in the Mayans of the Yucatan Penninsula, Mexico. *Clin. Genet.* **39**:26-32.

Kane, J.P, D.A. Hardman, and H.E. Paulus. 1980. Heterogeneity of apolipoprotein B: Isolation of a new species from human chylomicrons. *Proc. Natl. Acad. Sci. U. S. A.* **77**:2465-2469.

Karathanasis, S.K. 1985. Apolipoprotein multigene family: tandem organization of human apolipoprotein AI, CIII, and AIV genes. *Proc. Natl. Acad. Sci. U. S. A.* **82**:6374-6378.

Karathanasis, S.K., V.I. Zannis, and J.L. Breslow. 1983*a*. Isolation and characterization of the human apolipoprotein A-I gene. *Proc. Natl. Acad. Sci. U. S. A.* **80**:6147-6151.

Karathanasis, S.K., V.I. Zannis, and J.L. Breslow. 1983*b*. A DNA insertion in the apolipoprotein A-I gene of patients with premature atherosclerosis. *Nature* **305**:823-825.

Kawate, R., M. Yamakido, Y. Nishimoto, P.H. Bennett, R.F. Hamman, and W.C. Knowler. 1979. Diabetes mellitus and its complications in Japanese migrants on the island of Hawaii. *Diabetes Care* **2**:161-170.

Kayano, T., C.F. Burant, H. Fukumoto, G.W. Gould, Y. Fan, R.L. Eddy, M.G. Byers, T.B. Shows, S. Seino, and G.I. Bell. 1990. Human facilitative glucose transporters: isolation, functional characterization, and gene localization of cDNAs encoding an isoform (GLUT5) expressed in small intestine, kidney, muscle and adipose tissue and an unusual glucose transporter pseudogene-like sequence (GLUT6). *J. Biol. Chem.* **265**:13276-13282.

Kelly, R.B., N.R. Cozzarelli, M.P. Deutscher, I.R. Lehman, and A. Kornberg. 1970. Enzymatic synthesis of deoxyribonucleic acid. XXXII. Replication of duplex deoxyribonucleic acid by polymerase at a single strand break. *J. Biol. Chem.* **245**:39-45.

Kessling, A.M., B. Horsthemke, and S.E. Humphries. 1985. A study of DNA polymorphisms around the human apolipoprotein AI gene in hyperlipidaemic and normal individuals. *Clin. Genet.* **28**:296-306.

Kessling, A.M., K. Berg, E. Mockleby, and S.E. Humphries. 1986. DNA polymorphisms around the apo AI gene in normal and hyperlipidaemic individuals selected for a twin study. *Clin. Genet.* **29**:485-490.

Khoury, M.J., T.H. Beaty, and W.D. Flanders. 1990. Epidemiologic approaches to the use of DNA markers in the search for disease susceptibility genes. *Epidemiol. Rev.* **12**:41-55.

King, H., P. Zimmet, K. Pargeter, L.R. Raper, and V. Collins. 1984. Ethnic differences in susceptibility to non-insulin-dependent diabetes. A comparative study of two urbanized Micronesian populations. *Diabetes* **33**:1002-1007.

Kirchgessner, T.G., J.C. Chuat, C. Heinzmann, J. Etienne, S. Guilhot, K. Svenson, D. Ameis, C. Pilon, L. d'Auriol, A. Andalibi, M.C. Schotz, F. Galibert, and A.J. Lusis. 1989. Organization of the human lipoprotein lipase gene and evolution of the lipase gene family. *Proc. Natl. Acad. Sci. U. S. A.* **86**:9647-9651.

Kissebah, A.H., and A.N. Peiris. 1989. Biology of regional body fat distribution: relationship to non-insulin-dependent diabetes mellitus. *Diabetes Metab. Rev.* **5**:83-109.

Klasen, E.C., M. Smit, P. de Knijff, J.G. Leuven, R. Kempen-Voogd, and L. Havekes. 1987. Apolipoprotein E phenotype and gene distribution in the Netherlands. *Hum. Hered.* **37**:340-344.

Knott, T.J., R.L. Eddy, M.E. Robertson, L.M. Priestley, J. Scott, and T.B. Shows. 1984*a*. Chromosomal localization of the human apoprotein CI gene and of a polymorphic apoprotein AII gene. *Biochem. Biophys. Res. Commun.* **125**:299-306.

Knott, T.J., L.M. Priestley, M. Urdea, and J. Scott. 1984*b*. Isolation and characterisation of a cDNA encoding the precursor for human apolipoprotein AII. *Biochem. Biophys. Res. Commun.* **120**:734-740.

Knott, T.J., M.E. Robertson, L.M. Priestley, M. Urdea, S. Wallis, and J. Scott. 1984*c*. Characterisation of mRNAs encoding the precursor for human apolipoprotein CI. *Nucleic Acids Res.* **12**:3909-3915.

Knott, T.J., R.J. Pease, L.M. Powell, S.C. Wallis, S.C. Rall, B. Innerarity, B. Blackhart, W.H. Taylor, Y. Marcel, R. Milne, D. Johnson, M. Fuller, A.J. Lusis, B.J. McCarthy, R.W. Mahley, B. Levy-Wilson, and J. Scott. 1986. Complete protein sequence and identification of structural domains of human apolipoprotein B. *Nature* **323**:734-738.

Knowler, W.C., P.H. Bennett, R.F. Hamman, and M. Miller. 1978. Diabetic incidence and prevalence in Pima Indians: a 19-fold greater incidence than in Rochester, Minnesota. *Am. J. Epidemiol.* **108**:497-504.

Knowler, W.C., D.J. Pettitt, and R.G. Nelson. 1989. Ten-year prediction of diabetes by hyperinsulinemia and parental diabetes in Pima Indian children with normal glucose tolerance. *Diabetes* **38**:92A (Abstract).

Knowler, W.C., P.J. Pettit, M.F. Saad, and P.H. Bennett. 1990. Diabetes mellitus in Pima Indians: incidence, risk factors and pathogenesis. *Diabetes Metab. Rev.* **6**:1-27.

Kobberling, J., and H. Tillil. 1982. Empirical risk figures for first degree relatives of non-insulin dependent diabetics. In *The Genetics of Diabetes Mellitus*. J. Kobberling, and R. Tattersall, editors. Academic Press, London. pp. 201-209.

Kobori, S., N. Nakamura, H. Uzawa, and M. Shichiri. 1988. Influence of apolipoprotein E polymorphism on plasma lipid and apolipoprotein levels, and clinical characteristics of type III hyperlipoproteinemia due to apolipoprotein E phenotype E2/2 in Japan. *Atherosclerosis* **69**:81-88.

Kontula, K., K. Aalto-Setala, T. Kuusi, L. Hamalainen, and A.C. Syvanen. 1990. Apolipoprotein E polymorphism determined by restriction enzyme analysis of DNA amplified by polymerase chain reaction: convenient alternative to phenotyping by isoelectric focusing. *Clin. Chem.* **36**:2087-2092.

Kostner, G.M. 1974. Studies of the structure and composition of human serum lipoproteins: Isolation and partial characterization of apolipoprotein A-III. *Biochim. Biophys. Acta* **336**:383-395.

Kotze, M.J., E. Langenhoven, A.E. Retief, K. Steyn, M.P. Marais, J.J. Grobbelaar, C.J. Oosthuizen, H.F. Weich, and A.J. Benade. 1987. Haplotype associations of three DNA polymorphisms at the human low density lipoprotein receptor gene locus in familial hypercholesterolaemia. *J. Med. Genet.* **24**:750-755.

Kotze, M.J., E. Langenhoven, A.E. Retief, H.C. Seftel, H.E. Henderson, and H.F. Weich. 1989. Haplotypes identified by 10 DNA restriction fragment length polymorphisms at the human low density lipoprotein receptor gene locus. *J. Med. Genet.* **26**:255-259.

Kuzuya, T., and Japan Diabetes Society Committee on Diabetic Twins. 1988. Diabetes mellitus in twins: a cooperative study in Japan. *Diabetes Res. Clin. Pract.* **5**:271-280.

Lackner, K.J., S.W. Law, and H.B. Brewer Jr. 1985. The human apolipoprotein A-II gene: complete nucleic acid sequence and genomic organization. *Nucleic Acids Res.* **13**:4597-4608.

Lander, E., and D. Botstein. 1986*a*. Mapping complex genetic traits in humans: new methods using a complete RFLP linkage map. *Cold Spring Harb. Symp. Quant. Biol.* **51** (Pt 1):49-62.

Lander, E., and D. Botstein. 1986*b*. Strategies for studying heterogeneous genetic traits in humans by using a linkage map of restriction fragment length polymorphism. *Proc. Natl. Acad. Sci. U. S. A.* **83**:7353-7357.

Langlois, S., S. Deeb, J.D. Brunzell, J.J. Kastelein, and M.R. Hayden. 1989. A major insertion accounts for a significant proportion of mutations underlying human lipoprotein lipase deficiency. *Proc. Natl. Acad. Sci. U. S. A.* **86**:948-952.

Lapidus, L., C. Bengtsson, B. Larsson, K. Pennert, E. Rybo, and L. Sjostrom. 1984. Distribution of adipose tissue and risk of cardiovascular disease and death: a 12 year follow-up of participants in the population study of women in Gothenburg, Sweden. *Br. Med. J.* **289**:1257-1261.

Law, S.W., G. Gray, H.B. Brewer Jr., A.Y. Sakaguchi, and S.L. Naylor. 1984. Human apolipoprotein A-I and C-III genes reside in the p11-q13 region of chromosome 11. *Biochem. Biophys. Res. Commun.* **118**:934-942.

Law, A., S.C. Wallis, L.M. Powell, R.J. Pease, H. Brunt, L.M. Priestley, T.J. Knott, J. Scott, D.G. Altman, G.J. Miller, and N.E. Miller. 1986. Common DNA polymorphism within coding sequence of apolipoprotein B gene associated with altered lipid levels. *Lancet* **i**:1301-1303.

Lee, N.S., H.B. Brewer Jr., and J.C. Obsorne Jr. 1983. β2-glycoprotein I: molecular properties of an unusual apolipoprotein H. *J. Biol. Chem.* **258**:4765-4770.

Lehtimaki, T., T. Moilanen, J. Viikari, H.K. Akerblom, C. Ehnholm, T. Ronnemaa, J. Marniemi, G. Dahlen, and T. Nikkari. 1990. Apolipoprotein E phenotypes in Finnish youths: a cross-sectional and 6-year follow-up study. *J. Lipid Res.* **31**:487-495.

Lenzen, H.J., G. Assmann, R. Buchwalsky, and H. Schulte. 1986. Association of apolipoprotein E polymorphism, low-density lipoprotein cholesterol, and coronary artery disease. *Clin. Chem.* **32**:778-781.

Leppert, M.F., S.J. Hasstedt, T. Holm, P. O'Connell, L. Wu, O. Ash, R.R. Williams, and R. White. 1986. A DNA probe for the LDL receptor gene is tightly linked to hypercholesterolemia in a pedigree with early coronary disease. *Am. J. Hum. Genet.* **39**:300-306.

Leren, T.P., K. Berg, I. Hjermann, and P. Leren. 1988. Further evidence for an association between the *XbaI* polymorphism at the apolipoprotein B locus and lipoprotein level. *Clin. Genet.* **34**:347-351.

Li, C.C. 1955. *Population Genetics*. University of Chicago Press, Chicago. pp. 366.

Lillioja, S., D.M. Mott, J.K. Zawadski, A.A. Young, W.G.H. Abbott, W.C. Knowler, P.H. Bennett, P. Moll, and C. Bogardus. 1987. *In vivo* insulin action is a familial characteristic in non-diabetic Pima Indians. *Diabetes* **36**:1329-1335.

Lillioja, S., D.M. Mott, B.V. Howard, P.H. Bennett, H. Yki-Jarvinen, D. Freymond, B.L. Nyombo, F. Zurlo, B. Swinburn, and C. Bogardus. 1988. Impaired glucose tolerance as a disorder of insulin action: longitudinal and cross-sectional studies in Pima Indians. *N. Engl. J. Med.* **318**:1217-1225.

Lim, D., R.T. Coleman, G. Assmann, and P.M. Frossard. 1986. Deletion of an *Alu* sequence in the 5′ of the apolipoprotein AI associated with decreased serum HDL-cholesterl levels. *Am. J. Hum. Genet.* **39**:Suppl. pp. 621.(Abstract).

Lindgren, V., K.L. Luskey, D.W. Russell, and U. Francke. 1985. Human genes involved in cholesterol metabolism: Chromsomal mapping of the loci for the low density lipoprotein receptor and 3-hydroxy-3-methylglutaryl-coenzyme A reductase with cDNA probes. *Proc. Natl. Acad. Sci. U. S. A.* **82**:8567-8571.

Luo, C.C., W.H. Li, M.N. Moore, and L. Chan. 1986. Structure and evolution of the apolipoprotein multigene family. *J. Mol. Biol.* **187**:325-340.

Lusis, A.J., C. Heinzmann, R.S. Sparkes, J. Scott, T.J. Knott, R. Geller, M.C. Sparkes, and T. Mohandas. 1986. Regional mapping of human chromosome 19: organization of genes for plasma lipid transport (APOC1, -C2, and -E and LDLR) and the genes C3, PEPD, and GPI. *Proc. Natl. Acad. Sci. U. S. A.* **83**:3929-3933.

Malloy, M.J., J.P. Kane, D.A. Hardman, R.L. Hamilton, and K.B. Dolal. 1981. Normotriglyceridemic abetalipoproteinaemia: absence of the B-100 apolipoprotein. *J. Clin. Invest.* **67**:1441-1450.

Maniatis, T., F.F. Fritsch, and J. Sambrook. 1982. *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor Laboratory, New York. pp. 545.

Maruyama, T., and P.A. Fuerst. 1985. Population bottlenecks and nonequilibrium models in population genetics. II. Number of alleles in a small population that was formed by a recent bottleneck. *Genetics* **111**:675-689.

Mather, H.M., and H. Keen, 1985. The Southall diabetes survey prevalence of diabetes in Asians and Europeans. *Br. Med. J.* **291**:1081-1084.

Mazzone, T., D. Foster, and A. Chait. 1984. *In vivo* stimulation of low density lipoprotein degradation by insulin. *Diabetes* **33**:333-338.

Modan, M., A. Karasik, H. Halkin, Z. Fuchs, A. Lusky, A. Shitrit and B. Modan. 1986. Effect of past and concurrent body mass index on prevalence of glucose intolerance and type 2 (non-insulin dependent) diabetes and on insulin response; the Israeli study of glucose intolerance, obesity and hypertension. *Diabetologia* **29**:82-89.

Monaco, A.P., R.L. Neve, C. Colletti-Feener, C.J. Bertelson, D.M. Kurnit, and L.M. Kunkel. 1986. Isolation of candidate cDNAs for portions of the Duchenne muscular dystrophy gene. *Nature* **323**:646-650.

Monsalve, M.V., H. Henderson, G. Roederer, P. Julien, S. Deeb, J.J. Kastelein, L. Peritz, R. Devlin, T. Bruin, M.R. Murthy, C. Gagne, J. Davignon, P.J. Lupien, J.D. Brunzell, and M.R. Hayden. 1990. A missense mutation at codon 188 of the human lipoprotein lipase gene is a frequent cause of lipoprotein lipase deficiency in persons of different ancestries. *J. Clin. Invest.* **86**:728-734.

Morton, N.E. 1955. Sequential tests for the detection of linkage. *Am. J. Hum. Genet.* **7**:277-318.

Morton, N.E. 1956. The detection and estimation of linkage between elliptocytosis and Rh blood types. *Am. J. Hum. Genet.* **8**:80-96.

Motulsky, A.G. 1989. Genetic aspects of familial hypercholesterolemia and its diagnosis. *Arteriosclerosis* **9**:I3-I7.

Murai, A., T. Miyahara, N. Fujimoto, M. Matsuda, and M. Kameyama. 1986. Lp(a) lipoprotein as a risk factor for coronary heart disease and cerebral infarction. *Atherosclerosis* **59**:199-204.

Myklebost, O., S. Rogne, I. Hjermann, B. Olaisen, and H. Prydz. 1990. Association analysis of lipid levels and apolipoprotein restriction fragment length polymorphisms. *Hum. Genet.* **86**:209-214.

Nakaya, Y., J. Schaefer, and H.B. Brewer Jr. 1980. Activation of human post heparin lipoprotein lipase by apolipoprotein H (β2-glycoprotein I). *Biochem. Biophys. Res. Commun.* **95**:1168-1172.

Nasmyth, K.A. 1982. The regulation of yeast mating-type chromatin structure by SIR: an action at a distance affecting both transcription and transposition. *Cell* **30**:567-578.

National Diabetes Data Group. 1979. Classification and diagnosis of diabetes mellitus and other categories of glucose intolerance. *Diabetes* **28**:1039-1057.

Nikkila, E.A. 1978. Metabolic and endocrine control of plasma high density lipoprotein concentrations. In *High-density Lipoproteins and Atherosclerosis*. A.M. Gotto, N.E. Miller, and M.F. Oliver, editors. Elsevier Science Publishers, Amsterdam. pp. 177-192.

Nikkila, E.A., and P. Hormila. 1978. Serum lipids and lipoproteins in insulin-treated diabetes. *Diabetes* **27**:1078-1086.

Ohlson, L.O., B. Larsson, K. Svardsudd, L. Welin, H. Eriksson, L. Wilnelmsen, P. Bjorntorp, and G. Tibblin. 1985. The influence of body fat distribution on the incidence of diabetes mellitus. 13.5 years of follow-up of the participants in the study of men born in 1913. *Diabetes* **34**:1055-1058.

Olofsson, S.O., W.J. McConathy, and P. Alaupovic. 1978. Isolation and partial characterization of a new acidic apolipoprotein (apolipoprotein F) from high density lipoproteins of human plasma. *Biochemistry* **17**:1032-1036.

Olofsson, S.O., G. Bjursell, K. Bostrom, P. Carlsson, J. Elovson, A.A. Protter, M.A. Reuben, and G. Bondjers. 1987. Apolipoprotein B: structure, biosynthesis and role in the lipoprotein assembly process. *Atherosclerosis* **68**:1-17.

O'Rahilly, S., R.S. Spivey, R.R. Holman, Z. Nugent, A. Clark, and R.C. Turner. 1987. Type II diabetes of early onset: a distinct clinical and genetic syndrome? *Br. Med. J.* **294**:923-928.

Ordovas, J.M., L. Letwack-Klein, P.W.F. Wilson, M.M. Schaeffer, and E.J. Schaeffer. 1987. Apolipoprotein E isoform phenotyping methodology and population frequency with identification of apo E1 and E5 isoforms. *J. Lipid Res.* **28**:371-380.

Orita, M., Y. Suzuki, T. Sekiya, and K. Hayashi. 1989. Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. *Genomics* **5**:874-979.

Ott, J. 1974. Estimation of the recombination fraction in human pedigrees: efficient computation of the likelihood for human linkage studies. *Am. J. Hum. Genet.* **26**:588-597.

Parhofer, K.G., W.O. Richter, and P. Schwandt. 1990. Apolipoprotein E phenotype frequency in type II diabetic patients with different forms of hyperlipoproteinemia. *Horm. Metab. Res.* **22**:589-594.

Paul, H., D. Galton, and J. Stocks. 1987. DNA polymorphic patterns and haplotype arrangements of the apo A-1, apo C-III, apo A-IV gene cluster in different ethnic groups. *Hum. Genet.* **75**:264-268.

Paulweber, B., W. Friedl, F. Krempler, S.E. Humphries, and F. Sandhofer. 1988. Genetic variation in the apolipoprotein AI-CIII-AIV gene cluster and coronary heart disease. *Atherosclerosis* **73**:125-133.

Payne, C.D. (editor) 1986. *Generalized Linear Interactive Modelling Manual*. Nuffield Press, Oxford. pp. 305.

Pedersen, J.C., and K. Berg. 1990. Gene-gene interaction between the low density lipoprotein receptor and apolipoprotein E loci affects lipid levels. *Clin. Genet.* **38**:287-294.

Polz, E., and G.M. Kostner. 1979. The binding of β2-glycoprotein I to human serum lipoproteins: distribution among density fractions. *FEBS Lett.* **102**:183-186.

Price, W.H., S.W. Morris, R. Burgon, P.M. Donald, and A.H. Kitchin. 1986. Apolipoprotein CIII polymorphism and coronary heart disease [letter]. *Lancet* **ii**:1041.

Quarfordt, S.H., G. Michalopoulos, and B. Schirmer. 1982. The effect of human C apolipoproteins on the *in vitro* hepatic metabolism of triglyceride emulsions in the rat. *J. Biol. Chem.* **257**:14642-14647.

Rajput-Williams, J., J. Eyre, M.N. Nanjee, D. Crook, J. Scott, and N.E. Miller. 1989. Plasma lipoprotein lipids in relation to the *Msp*I polymorphism of the apolipoprotein AII gene in Caucasian men. Lack of association with plasma triglyceride concentration. *Atherosclerosis* **77**:31-36.

Rall, S.C.Jr., K.H. Weisgraber, and R.W. Mahley. 1982. Human apolipoprotein E. The complete amino acid sequence. *J. Biol. Chem.* **257**:4171-4178.

Reaven, G.M., and R. Miller. 1968. Study of the relationship between glucose and insulin reponses to an oral glucose load in man. *Diabetes* **17**:560-569.

Reaven, G.M., R. Bernstein, B. Davis, and J.M. Olefsky. 1976. Non-ketotic diabetes mellitus: insulin deficiency or insulin resistance. *Am. J. Med.* **60**:80-88.

Reckless, J.P.D., D.J. Betteridge, P. Wu, B. Payne, and D.J. Galton. 1978. High and low density lipoproteins and the prevalence of vascular disease in diabetes mellitus. *Br. Med. J.* **1**:883-886.

Reed, K.C., and D.A. Mann. 1985. Rapid transfer of DNA from agarose gels to nylon membranes. *Nucleic Acids Res.* **13**:7207-7221.

Reeders, S.T., M.H. Breuning, K.E. Davies, R.D. Nicholls, A.P. Jarmam, D.R. Higgs, P.L. Pearson, and D.J. Weatherall. 1985. A highly polymorphic DNA marker linked to adult polycystic kidney disease on chromosome 16. *Nature* **317**:542-544.

Rees, A., C.C. Shoulders, J. Stocks, D.J. Galton, and F.E. Baralle. 1983. DNA polymorphism adjacent to human apoprotein A-1 gene: relation to hypertriglyceridaemia. *Lancet* **i**:444-446.

Rees, A., J. Stocks, C.R. Sharpe, M.A. Vella, C.C. Shoulders, J. Katz, N.I. Jowett, F.E. Baralle, and D.J. Galton. 1985. Deoxyribonucleic acid polymorphism in the apolipoprotein A-1-C-III gene cluster. Association with hypertriglyceridemia. *J. Clin. Invest.* **76**:1090-1095.

Rees, A., J. Stocks, H. Paul, Y. Ohuchi, and D. Galton. 1986. Haplotypes identified by DNA polymorphisms at the apolipoprotein A-1 and C-III loci and hypertriglyceridaemia. A study in a Japanese population. *Hum. Genet.* **72**:168-171.

Robertson, F.W., and A.M. Cumming. 1985. Effects of apoprotein E polymorphism on serum lipoprotein concentration. *Arteriosclerosis* **5**:283-292.

Rothman, K.J. 1990. No adjustments are needed for multiple comparisons. *Epidemiology* **1**:43-46.

Rushforth, N.B., P.H. Bennett, A.G. Sternberg, T.A. Burch, and M. Miller. 1971. Diabetes in Pima Indians: evidence of bimodality in glucose tolerance distributions. *Diabetes* **20**:756-765.

Russell, D.W., M.A. Lehrman, T.C. Sudhof, T. Yamamoto, C.G. Davis, H.H. Hobbs, M.S. Brown, and J.L. Goldstein. 1986. The LDL receptor in familial hypercholesterolemia: use of human mutations to dissect a membrane protein. *Cold Spring Harb. Symp. Quant. Biol.* **51** (Pt 2):811-819.

Saad, M.F., W.C. Knowler, D.J. Pettitt, R.G. Nelson, D.M. Mott, and P.H. Bennett. 1988. The natural history of impaired glucose tolerance in the Pima Indians. *N. Engl. J. Med.* **319**:1500-1506.

Saiki, R.K., S. Scharf, F. Faloona, K.B. Mullis, G.T. Horn, H.A. Erlich, and N. Arnheim. 1985. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **230**:1350-1354.

Savolainen, M.J., T. Korhonen, K. Aalto-Setala, K. Kontula, and Y.A. Kesaniemi. 1991. Screening patients for a prevalent LDL receptor mutation in patients with severe hypercholesterolaemia. *Hum. Genet.* **87**:125-128.

Schleicher, E., B. Olgemoller, J. Schon, T. Durst, and O.H. Weiland. 1985. Limited non-enzymatic glucosylation of low density lipoprotein does not alter its catabolism in tissue culture. *Biochim. Biophys. Acta* **846**:226-233.

Schonfeld, G., C. Birge, J.P. Miller, G. Kessler, and J. Santiago. 1974. Apolipoprotein B levels and altered lipoprotein composition in diabetes. *Diabetes* **23**:827-834.

Schooneveldt, M., T. Songer, P. Zimmet, and K. Thoma. 1988. Changing mortality patterns in Nauruans - an example of an epidemiological transition. *J. Epidemiol. Community Health* **42**:89-95.

Schull, W.J., and K.M. Weiss. 1980. Genetic epidemiology: four strategies. *Epidemiol. Rev.* **2**:1-18.

Schuster, H., B. Stiefenhofer, G. Wolfram, C. Keller, S. Humphries, A. Huber, and N. Zollner. 1989. Four DNA polymorphisms in the LDL-receptor gene and their use in diagnosis of familial hypercholesterolemia. *Hum. Genet.* **82**:69-72.

Schuster, H., S. Humphries, G. Ranh, C. Held, C. Keller, G. Wolfman, and N. Zollner. 1990. Association of DNA-haplotypes in the human LDLR-receptor gene with normal serum cholesterol levels. *Clin. Genet.* **38**:401-409.

Scott, J., T.J. Knott, L.M. Priestley, M.E. Robertson, D.V. Mann, G. Kostner, G.J. Miller, and N.E. Miller. 1985. High-density lipoprotein composition is altered by a common DNA polymorphism adjacent to apoprotein AII gene in man. *Lancet* **i**:771-773.

Seilhamer, J.J., A.A. Protter, P. Frossard, and B. Levy Wilson. 1984. Isolation and DNA sequence of full-length cDNA and of the entire gene for human apolipoprotein AI - discovery of a new genetic polymorphism in the apo AI gene. *DNA* **3**:309-317.

Sepehrnia, B., M.I. Kamboh, L.L. Adams Campbell, M. Nwankwo, and R.E. Ferrell. 1988*a*. Genetic studies of human apolipoproteins. VII. Population distribution of polymorphisms of apolipoproteins A-I, A-II, A-IV, C-II, E, and H in Nigeria. *Am. J. Hum. Genet.* **43**:847-853.

Sepehrnia, B., M.I. Kamboh, and R.E. Ferrell. 1988*b*. Genetic studies of human apolipoproteins. III. Polymorphism of apolipoprotein C-II. *Hum. Hered.* **38**:136-143.

Serjeantson, S. and P. Zimmet. 1984. Diabetes in the Pacific: evidence for a major gene. In *Diabetes Mellitus; Recent Knowledge on Aetiology, Complications and Treatment*. S. Baba, M. Gould, and P. Zimmet, editors. Academic Press, Sydney. pp. 23-30.

Serjeantson, S.W., and P. Zimmet. 1989. Genetics of NIDDM: pilgrim's progress. In *Frontiers of Diabetes Research: Current Trends in Non-insulin Dependent Diabetes mellitus*. K.G.M.M. Alberti, and R. Mazze, editors. Elsevier Science Publishers, Amsterdam. pp. 21-35.

Serjeantson, S.W., and P. Zimmet. 1991. Genetics of non-insulin dependent diabetes mellitus in 1990. *Baillere's Clin. Endocrin. Metab.* **5**:477-493.

Serjeantson, S.W., D. Owerbach, P. Zimmet, J. Nerup, and K. Thoma. 1983. Genetics of diabetes in Nauru: effects of foreign admixture, HLA antigens and the insulin-gene-linked polymorphism. *Diabetologia* **25**:13-17.

Serjeantson, S.W., B. White, G.I. Bell, and P. Zimmet. 1987. RFLPs in the insulin receptor gene and type 2 diabetes in the Pacific. In *New Approaches to Genetic Diseases*. T, Sasasuki, editor. Academic Press, London. pp. 23-28.

Serjeantson, S.W., B. White, G.I Bell, and P. Zimmet. 1989. The glucose transporter gene and type 2 diabetes in the Pacific. In *Diabetes*. Proceedings of the 13th International Diabetes Federation Congress. R.G. Larkins, P.Z. Zimmet, and D.J. Chisholm, editors. Excerpta Medica, Amsterdam. pp. 329-333.

Sharp, P.A., B. Sugden, and J. Sambrook. 1973. Detection of two restriction endonuclease activities in *Haemophilus parainfluenzae* using analytical agarose-ethidium bromide electrophoresis. *Biochemistry* **12**:3055-3063.

Sheffield, V.C., D.R. Cox, L.S. Lerman, and R.M. Myers. 1989. Attachment of a 40-base-pair G + C-rich sequence (GC-clamp) to genomic DNA fragments by the polymerase chain reaction results in improved detection of single-base changes. *Proc. Natl. Acad. Sci. U. S. A.* **86**:232-236.

Shows, T.B., R.L. Eddy, M.G. Byers, Y. Fukushima, C.R. Dehaven, J.C. Murray, and G.I. Bell. 1987. Polymorphic human glucose transporter gene (GLUT) is on chromosome 1p31.3-p35. *Diabetes* **36**:546-549.

Shulman, R.S., P.N. Herbert, K. Wehrly, and D.S. Fredrickson. 1975. The complete amino acid sequence of CI (apoLp-Ser), an apolipoprotein from very low density lipoprotein. *J. Biol. Chem.* **250**:182-190.

Sicree, R.A., P.Z. Zimmet, H.O. King, and J.S. Coventry. 1987. Plasma insulin response among Nauruans. Prediction of deterioration in glucose tolerance over 6 years. *Diabetes* **36**:179-186.

Simpson, T.B. 1844. Pacific navigation and British seamen. *Nautical Magazine and Naval Chronicle* **13**:102.

Simpson, I.A., and S.W. Cushman. 1986. Hormonal regulation of glucose transport. *Annu. Rev. Biochem.* **55**:1059-1089.

Simpson, N.E., K.K. Kidd, P.J. Goodfellow, H. McDermid, S. Myers, J.R. Kidd, C.E. Jackson, A.M.V. Duncan, L.A. Farrer, K. Brasch, C. Castiglione, M. Genel, J. Gertner, C.R. Greenberg, J.F. Gusella, J.J.A. Holden, and B.N. White. 1987. Assignment of multiple endocrine neoplasia Type 2A to chromosome 10 by linkage analysis. *Nature* **328**:528-530.

Sing, C.F., and J. Davignon. 1985. Role of the apolipoprotein E polymorphism in determining normal plasma lipid and lipoprotein variation. *Am. J. Hum. Genet.* **37**:268-285.

Sivitz, W.I., S.L. de Sautel, T. Kayano, G.I. Bell, and J.E. Pessin. 1989. Regulation of glucose transporter messenger RNA in insulin-deficient states. *Nature* **340**:72-74.

Smeets, H.J., J. Poddighe, P.M. Stuyt, A.F. Stalenhoef, H.H. Ropers, and B. Wieringa. 1988. Identification of apolipoprotein E polymorphism by using synthetic oligonucleotides. *J. Lipid Res.* **29**:1231-1237.

Smit, M., E. van der Kooij-Meijs, R.R. Frants, L. Havekes, and E.C. Klasen. 1988. Apolipoprotein gene cluster on chromosome 19. Definite localization of the APOC2 gene and the polymorphic *Hpa*I site associated with type III hyperlipoproteinemia. *Hum. Genet.* **78**:90-93.

Smith, L.C., H.J. Pownall, and A.M. Gotto Jr. 1978. The plasma lipoproteins: Structure and metabolism. *Ann. Rev. Biochem.* **47**:751-779.

Snedecor, G.W., and W.G. Cochran. 1980. *Statistical Methods*. University Press, Ames. pp. 505.

Snowden, C., R.S. Houlston, M.H. Arif, M.F. Laker, S.E. Humphries, and K.G.M.M. Alberti. 1991. Disparity between apolipoprotein E phenotypes and genotypes (as determined by polymerase chain reaction and oligonucleotide probes) in patients with non-insulin-dependent diabetes mellitus. *Clin. Chim. Acta* **196**:49-58.

Soutar, A.K., C.W. Garner, H.N Baker, J.T. Sparrow, R.L. Jackson, A.M. Gotto Jr., and L.C. Smith. 1975. Effect of the human apolipoproteins and phosphatidylcholine acyl donor on the activity of lecithin:cholesterol acyltransferase. *Biochemistry* **14**:3057-3064.

Southern, E.M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**:503-517.

SPSS Inc. 1986. *SPSSx User's Guide*. McGraw-Hill, New York. pp. 988.

Stanhope, J.M., and I.A.M. Prior. 1980. The Tokelau Island migrant study: prevalence and incidence of diabetes mellitus. *NZ Med. J.* **92**:417-421.

Steinbrecher, U., and J.L. Witztum. 1984. Glucosylation of low density lipoproteins to an extent comparable to that seen in diabetics slows their catabolism. *Diabetes* **33**:130-134.

Steyrer, E., and G.M. Kostner. 1988. Activation of lecithin-cholesterol acyltransferase by apolipoprotein D: comparison of proteoliposomes containing apolipoprotein D, A-I or C-I. *Biochim. Biophys. Acta* **958**:484-491.

Stocks, J., H. Paul, and D. Galton. 1987. Haplotypes identified by DNA restriction-fragment-length polymorphisms in the A-1 C-III A-IV gene region and hypertriglyceridemia. *Am. J. Hum. Genet.* **41**:106-118.

Sudhof, T.C., J.L. Goldstein, M.S. Brown, and D.W. Russell. 1985. The LDL receptor gene: a mosaic of exons shared with different proteins. *Science* **228**:815-822.

Suzuku, K., and T. Kono. 1980. Evidence that insulin causes translocation of glucose transport activity to the plasma membrane from an intracellular storage site. *Proc. Natl. Acad. Sci. U. S. A.* **77**:2542-2545.

Swofford, D.L., and R.B. Selander. 1981. *A Computer Program for the Analysis of Allelic Variation in Genetics: Users Manual*. Dept. Genetics and Development, University of Illinois, Urbana, IL, USA. pp. 71.

Syvanen, A.C., K. Aalto-Setala, L. Harju, K. Kontula, and H. Soderlund. 1990. A primer-guided nucleotide incorporation assay in the genotyping of apolipoprotein E. *Genomics* **8**:684-692.

Tairi, M., M. Tairi, N. Hashimoto, F. Shimada, Y. Suzuki, A. Kanatsuka, F. Nakamura, Y. Ebina, M. Tatibana, H. Makino, and S. Yoshida. 1989. Human diabetes associated with a deletion of the tyrosine kinase domain of the insulin receptor. *Science* **245**:63-66.

Tartof, K.D., and C.A. Hobbs. 1987. Improved media for growing plasmid and cosmid clones. *Focus* **9**:12.

Taylor, S.E., T. Kadowaki, H. Kadowaki, D. Accili, A. Cama, and C. McKeon. Mutations in insulin-receptor gene in insulin-resistant patients. *Diabetes Care* **13**:257-279.

Terry, R.B., P.D. Wood, W.L. Haskell, M.L. Stefanik, and R.M. Krauss. 1989. Regional adiposity patterns in relation to lipids, lipoprotein cholesterol, and lipoprotein subfraction mass in men. *J. Clin. Endocrinol. Metab.* **68**:191-199.

Thompson, E.A., S. Deeb, D. Walker, and A.G. Motulsky. 1988. The detection of linkage disequilibrium between closely linked markers: RFLPs at the AI-CIII apolipoprotein genes. *Am. J. Hum. Genet.* **42**:113-124.

Thorn, J.A., J.C. Chamberlain, J.C. Alcolado, K. Oka, L. Chan, J. Stocks, and D.J. Galton. 1990. Lipoprotein and hepatic lipase gene variants in coronary atherosclerosis. *Atherosclerosis* **85**:55-60.

Trembath, R.C., D.J. Thomas, T.J. Hendra, J.S. Yudkin, and D.J. Galton. 1987. Deoxyribonucleic acid polymorphism of the apoprotein AI-CIII-AIV gene cluster and coronary heart disease in non-insulin-dependent diabetes. *Br. Med. J. Clin. Res.* **294**:1577-1578.

Tsao, Y.K., C.F. Wei, D.L. Robberson, A.M. Gotto Jr., and L. Chan. 1985. Isolation and characterization of the human apolipoprotein A-II gene. Electron microscopic analysis of RNA:DNA hybrids, nucleotide sequence, identification of a polymorphic *Msp*I site, and general structural organization of apolipoprotein genes. *J. Biol. Chem.* **260**:15222-15231.

Utermann, G. 1985. Genetic polymorphism of apolipoprotein E - impact on plasma lipoprotein metabolism. In *Diabetes, Obesity and Hyperlipidemias-III*. G. Crepaldi, A. Tiengo, and G. Baggio, editors. Elsevier Science Publishers, Amsterdam. pp. 1-28.

Utermann, G. 1987. Apolipoprotein E polymorphism in health and disease. *Am. Heart J.* **113**:433-440.

Utermann, G., K.H. Vogelberg, A. Steinmetz, W. Schoenborn, N. Pruin, M. Jaeschke, M. Hess, and H. Canzler. 1979. Polymorphism of apolipoprotein E: Genetics of hyperlipoproteinemia type III. *Clin. Genet.* **15**:37-62.

Utermann, G., A. Steinmetz, and W. Weber. 1982. Genetic control of human apolipoprotein E polymorphism: comparison of one and two dimensional techniques of isoprotein analysis. *Hum. Genet.* **60**:344-351.

Utermann, G., I. Kindermann, H. Kaffarnik, and A. Steinmetz. 1984. Apolipoprotein E phenotypes and hyperlipidemia. *Hum. Genet.* **65**:232-236.

van Gaal, L., E. Rillaerts, W. Creten, and I. de Leeuw. 1988. Relationship of body fat distribution pattern to atherogenic risk factors in NIDDM. Preliminary results. *Diabetes Care* **11**:103-106.

Vella, M., A. Kessling, N. Jowett, A. Rees, J. Stocks, S. Wallis, and D. Galton. 1985. DNA polymorphisms flanking the apo A-1 and insulin genes and type III hyperlipidaemia. *Hum. Genet.* **69**:275-276.

Viviani, N. 1970. *Nauru: Phosphate and Political Progress.* Australian National University, Canberra. pp. 215.

Vogelberg, K.H., and E. Maucy. 1988. Apo E2 phenotypes in type II diabetics with and without insulin therapy. *Klin. Wochenschr.* **66**:690-693.

von Eckardstein, A., H. Funke, M. Walter, K. Altland, A. Benninghoven, and G. Assmann. 1990. Structural analysis of human apolipoprotein A-I variants. Amino acid substitutions are non-randomly distributed throughout the apolipoprotein A-I primary structure. *J. Biol. Chem.* **265**:8610-8617.

Wald, A. 1947. *Sequential Analysis.* John Wiley, New York. pp. 212.

Wardell, M.R., P.A. Suckling, and E.D. Janus. 1982. Genetic variation in human apolipoprotein E. *J. Lipid Res.* **23**:1174-1182.

Wei, C-F., Y-K. Tsao, D.L. Robberson, A.M. Gotto Jr., K. Brown, and L. Chan. 1985. The structure of the human apolipoprotein C-II gene. *J. Biol. Chem.* **260**:15211-15221.

Weisgraber, K.H., Y.M. Newhouse, and R.W. Mahley. 1988. Apolipoprotein E genotyping using the polymerase chain reaction and allele-specific oligonucleotide probes. *Biochem. Biophys. Res. Commun.* **157**:1212-1217.

Welborn, T.A., D.H. Curnow, J.T. Wearne, K.J. Cullen, M.G. McCall, and N.S. Stenhouse. 1968. Diabetes detected by blood-sugar measurement after a glucose load: report from Busselton survey. *Med. J. Aust.* **2**:778-783.

Wenham, P.R., C.R. Newton, and W.H. Price. 1991. Analysis of apolipoprotein E genotypes by the Amplification Refractory Mutation System. *Clin. Chem.* **37**:241-244.

White, R., and J.M. Lalouel. 1987. Investigation of genetic linkage in human families. *Adv. Hum. Genet.* **16**:121-228.

WHO Expert Committee on Diabetes Mellitus. 1980. *Second report*. World Health Organization, Geneva, (*Tech. Rep. Ser.* 646).

WHO Study Group. 1982. *Diabetes Mellitus*. World Health Organization, Geneva, (*Tech. Rep. Ser.* 727).

Williams, L.G., N.I. Jowett, M.A. Vella, S. Humphries, and D.J. Galton. 1985. Allelic variation adjacent to the human insulin and apolipoprotein C-II genes in different ethnic groups. *Hum. Genet.* **71**:227-230.

Wilson, C., M.R. Wardell, K.H. Weisgraber, R.W. Mahley, and D.A. Agard. 1991. Three-dimensional structure of the LDL receptor-binding domain of human apolipoprotein E. *Science* **252**:1817-1822.

Wion, K.L., T.G. Kirchgessner, A.J. Lusis, M.C. Schotz, and R.M. Lawn. 1987. Human lipoprotein lipase complementary DNA sequence. *Science* **235**:1638-1641.

Wojciechowski, A.P., M. Farrall, P. Cullen, T.M.E. Wilson, J.D. Bayliss, B. Farren, B.A. Griffin, M.J. Caslake, C.J. Packard, J. Sheperd, R. Thakker, and J. Scott. 1991. Familial combined hyperlipidaemia linked to the apolipoprotein AI-CIII-AIV gene cluster on chromosome 11q23-q24. *Nature* **349**:161-164.

Xiang, K.S., N.J. Cox, N. Sanz, P. Huang, J.H. Karam, and G.I. Bell. 1989. Insulin-receptor and apolipoprotein genes contribute to development of NIDDM in Chinese Americans. *Diabetes* **38**:17-23.

Yamamoto, T., C.G. Davis, M.S. Brown, W.J. Schneider, M.L. Casey, J.L. Goldstein, and D.W. Russell. 1984. The human LDL receptor: a cysteine-rich protein with multiple *Alu* sequences in its mRNA. *Cell* **39**:27-38.

Yamamura, T., H. Dudo, K. Ishikawa, and A. Yamamoto. 1979. Familial type I hyperlipoproteinaemia caused by C-II deficiency. *Atherosclerosis* **34**:53-57.

Yanisch-Perron, C., J. Vieira, and J. Messing. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**:103-119.

Zannis, V.I., and J.L. Breslow. 1981. Human very-low-density lipoprotein apolipoprotein E isoprotein polymorphism is explained by genetic modification and posttranslational modification. *Biochemistry* **20**:1033-1041.

Zannis, V.I., and J.L. Breslow. 1984. Genetic mutations affecting human lipoprotein metabolism. *Adv. Hum. Genet.* **14**:125-213.

Zannis, V.I., S.K. Karathanasis, H.T. Keutmann, G. Goldberger, and J.L. Breslow. 1983. Intracellular and extracellular processing of human apolipoprotein A-I: secreted apolipoprotein A-I isoprotein 2 is a propeptide. *Proc. Natl. Acad. Sci. U. S. A.* **80**:2574-2578.

Zannis, V.I., J. McPherson, G. Goldberger, S.K. Karathanasis, and J.L. Breslow. 1984. Synthesis, intracellular processing, and signal peptide of human apolipoprotein E. *J. Biol. Chem.* **259**:5495-5499.

Zimmet, P. 1982. Type 2 (non-insulin-dependent) diabetes - an epidemiological overview. *Diabetologia* **22**:399-411.

Zimmet, P., and S. Whitehouse. 1978. Bimodality of fasting and two-hour glucose tolerance distributions in a Micronesian population. *Diabetes* **27**:793-800.

Zimmet, P., P. Taft, A. Guinea, W. Guthrie, and K. Thoma. 1977. The high prevalence of diabetes mellitus on a Central Pacific Island. *Diabetologia* **13**:111-115.

Zimmet, P., H. King, R. Taylor, L.R. Raper, B. Balkau, J. Borger, W. Heriot, and K. Thoma. 1984. The high prevalence of diabetes mellitus, impaired glucose tolerance and diabetic retinopathy in Nauru - the 1982 survey. *Diabetes. Res.* **1**:13-18.

Zimmet, P., G. Dowse, R. La Porte, C. Finch, and C. Moy. 1989. Epidemiology - its contribution to understanding of the etiology, pathogenesis and prevention of diabetes mellitus. In *Diabetes Mellitus: Pathophysiology and Treatment*. W. Creutzfeld, and P. Lefebvre, editors. Springer Verlag, Berlin. pp. 5-26.

Zimmet, P., G. Dowse, C. Finch, S. Serjeantson, and H. King. 1990. The epidemiology and natural history of NIDDM-lessons from the South Pacific. *Diabetes. Metab. Rev.* **6**:91-124.